

1. FELADAT

K: Az eltolt lineáris preferenciával növekedő hálózatban egy k fokszámú csúcshoz való csatlódás valószínűsége $(k + \lambda) / \sum_l (l + \lambda) N_l$, ahol N_l a l fokszámú csúcsok száma. Az előadáson megmutattuk, hogy $\lambda = 0$ (lineáris preferencia) esetén a fokszámeloszlás nagy- k -ra hatvány alakú

$$P_k \sim k^{-3}$$

Vigyünk végig a lineáris preferenciára alkalmazott számolást a $\lambda = 1$ -re és mutassuk meg, hogy ekkor az eloszlás nagy- k alakja

$$P_k \sim k^{-4}$$

A fenti két eredmény azt sugallja, hogy (mint azt be is lehet bizonyítani) tetszőleges $\lambda > 0$ -ra a következő eredmény igaz

$$P_k \sim k^{-(k+\lambda)}$$

MEGOLDÁS

A Barabási–Albert-modell esetén $w_k = \frac{k}{\sum_l l N_l} = \frac{k}{A}$ rátával számoltunk, mellyel a Master-egyenletek a következő alakúak voltak:

$$\frac{dN_k}{dN} = -\frac{k}{A} N_k + \frac{k-1}{A} N_{k-1} \quad (1)$$

$$\frac{dN_1}{dN} = -\frac{1}{A} N_1 + 1 \quad (2)$$

Általánosan ezeket az egyenleteket $w_k = \frac{k+\lambda}{\sum_l (l+\lambda) N_l} = \frac{k+\lambda}{A_\lambda}$ esetre a következő formában írhatjuk fel:

$$\frac{dN_k}{dN} = -\frac{k+\lambda}{A_\lambda} N_k + \frac{k+\lambda-1}{A_\lambda} N_{k-1} \quad (3)$$

$$\frac{dN_1}{dN} = -\frac{1}{A_\lambda} N_1 + 1 \quad (4)$$

A gyakorlaton ismerttetettek alapján tudjuk, hogy az A_λ érték a következő módon egyszerűsíthető:

$$A_\lambda = \sum_l (l + \lambda) N_l = (2 + \lambda) N \quad (5)$$

Ezt helyettesítsük be a fenti (3)-as és (4)-es egyenletekbe:

$$\frac{dN_k}{dN} = -\frac{k+\lambda}{(2+\lambda)}N_k + \frac{k+\lambda-1}{(2+\lambda)}N_{k-1} \quad (6)$$

$$\frac{dN_1}{dN} = -\frac{1}{(2+\lambda)}N_1 + 1 \quad (7)$$

Definiáljuk a P_k élszámeloszlást az alábbi módon:

$$P_k = \frac{N_k}{N} \quad (8)$$

Ezt újból visszahelyettesítve az előző egyenletebe:

$$\frac{dN_k}{dN} = -\frac{k+\lambda}{(2+\lambda)}P_k + \frac{k+\lambda-1}{(2+\lambda)}P_{k-1} \quad (9)$$

$$\frac{dN_1}{dN} = -\frac{1}{(2+\lambda)}P_1 + 1 \quad (10)$$

Melyet szintén a (8)-as egyenlet átalakításából kapott $N_k = NP_k$ összefüggésből továbbvezetve:

$$\frac{d(NP_k)}{dN} = -\frac{k+\lambda}{(2+\lambda)}P_k + \frac{k+\lambda-1}{(2+\lambda)}P_{k-1} \quad (11)$$

$$\frac{d(NP_1)}{dN} = -\frac{1}{(2+\lambda)}P_1 + 1 \quad (12)$$

Elvégezve a bal oldalon álló deriváltakat:

$$P_k + N\frac{dP_k}{dN} = -\frac{k+\lambda}{(2+\lambda)}P_k + \frac{k+\lambda-1}{(2+\lambda)}P_{k-1} \quad (13)$$

$$P_1 + N\frac{dP_1}{dN} = -\frac{1}{(2+\lambda)}P_1 + 1 \quad (14)$$

A megmaradt P_k és P_1 értékeket átvisszük a jobb oldalra:

$$N\frac{dP_k}{dN} = -\frac{k+2+\lambda}{(2+\lambda)}P_k + \frac{k+\lambda-1}{(2+\lambda)}P_{k-1} \quad (15)$$

$$N\frac{dP_1}{dN} = -\frac{2+\lambda}{(2+\lambda)}P_1 + 1 \quad (16)$$

A feladat alapján a $\lambda = 1$ esetre kell meghatároznunk az élszámeloszlás értéket. Helyettesítjük be ez az értéket a kapott egyenletekbe:

$$N\frac{dP_k}{dN} = -\frac{k+2+1}{(2+1)}P_k + \frac{k+1-1}{(2+1)}P_{k-1} = -\frac{k+3}{3}P_k + \frac{k}{3}P_{k-1} \quad (17)$$

$$N \frac{dP_1}{dN} = -\frac{3+1}{(2+1)}P_1 + 1 = -\frac{4}{3}P_1 + 1 \quad (18)$$

Keressük a $P_k^{(st.)}$ stacionárius megoldást, mely esetében a deriváltak értéke 0:

$$0 = -\frac{k+4}{3}P_k^{(st.)} + \frac{k}{3}P_{k-1}^{(st.)} \quad (19)$$

$$0 = -\frac{4}{3}P_1^{(st.)} + 1 \quad (20)$$

A (20)-as egyenletből egyértelmű a $P_1^{(st.)}$ értéke:

$$P_1^{(st.)} = \frac{3}{4} \quad (21)$$

Míg a (19)-es egyenletből kiszámíthatjuk a $P_k^{(st.)}$ értéket is néhány átalakítás segítségével:

$$P_k^{(st.)} = \frac{k}{k+4}P_{k-1}^{(st.)} = \frac{k(k-1)}{(k+4)(k+3)}P_{k-2}^{(st.)} = \frac{k(k-1) \cdot \dots \cdot 1}{(k+4)(k+3) \cdot \dots \cdot 5}P_1^{(st.)} \quad (22)$$

Melybe behelyettesíthetjük a $P_1^{(st.)}$ értékét, majd egyszerűsíthetjük az egyenletet az azonos tagok kiejtésével a törtben:

$$P_k^{(st.)} = \frac{k(k-1) \cdot \dots \cdot 1}{(k+4)(k+3) \cdot \dots \cdot 5} \cdot \frac{3}{4} = \frac{1}{(k+4)(k+3)(k+2)(k+1)} \cdot (4 \cdot 3 \cdot 2 \cdot 1) \cdot \frac{3}{4} \quad (23)$$

Összevonva a maradékot a következő végeredményt kapjuk:

$$P_k^{(st.)} = \frac{18}{(k+4)(k+3)(k+2)(k+1)} \quad (24)$$

Ha k nagyon nagy, a mellette álló tagok elhagyhatóak:

$$P_k^{(st.)} = \frac{18}{k^4} \quad (25)$$

Tehát bebizonyítottuk a kezdeti feltevésünket, miszerint

$$\boxed{P_k \sim k^{-4}} \quad (26)$$

2. FELADAT

K: Az előadáson vizsgált lineáris preferenciával növekedő hálózatban egy k fokszámú csúcshoz való csatolódás valószínűsége $k / \sum_l l N_l$, ahol N_l az l fokszámú csúcsok száma. Az előadáson megmutattuk, hogy a fokszámeloszlás nagy- k -ra hatvány alakú $P_k \sim k^{-3}$. Feladatok:

- (i) Vizsgáljuk meg, hogy a fenti eredmény függ-e a kezdeti feltételektől! Indítsunk szimulációkat
 - (i.1) egy csúcsból,
 - (i.2) öt lineárisan csatolt csúcsból [$N_1(t=0) = 2, N_2(0) = 3, N_{k \neq 1,2} = 0$],
 - (i.3) öt kereszt alakban összekapcsolt csúcsból [$N_1(0) = 4, N_4(0) = 1, N_{k \neq 1,4} = 0$],s hasonlítsuk össze a fokszámeloszlások nagy- k -s viselkedését.
- (ii) Találjuk meg a maximális fokszámú csúcsot a fenti szimulációkban generált hálózatokban, s határozzuk meg a maximális fokszám átlagát $\langle k_{max} \rangle$ elég nagy csúcsszám esetére. Függ-e $\langle k_{max} \rangle$ a kezdeti feltételektől?

MEGOLDÁS

2.1. Megvalósítás

A feladatot megvalósító program forráskódját egy Jupyter Notebook-ban futó Python 3.7 kernel alatt írtam meg, valamint az ábrákat is ebben készítettem. Ezeket GitHub-on mind elérhetővé is tettem[1].

A második feladat során a Lineáris Preferenciális Modellt kellett szimuláljuk, tetszőleges N , vagy E csúcs/élszám darabszámra. A gráf szimulációjának működése az előző házi feladatban is ismertetett anti-preferenciális modellhez nagyon hasonló és a következő lépésekből áll:

- (1) Véletlenszerűen kiválasztunk egyet a már meglévő N csúcs közül.
- (2) Leszámoljuk a kiválasztott csúcs éleinek számát. Legyen ez a szám k .
- (3) A kiválasztott csúcshoz $w_k = k/A$ valószínűséggel csatolunk egy új csúcsot. Itt A a normalizációs állandó $A = \sum_l l N_l$, ahol N_l az l éllel rendelkező csúcsok száma.

Szintén az előző házi feladathoz hasonlóan a szimuláció futtatása összesen 3 darab adatsorral tért vissza. Ezek egy bemeneti paraméterrel választható módon a feladatban kirótt kezdeti feltételekkel voltak inicializálva, majd ezek minden lépésben frissültek. Végül a végleges értékeket használtam fel a további számításokban és ábrázolásokban. Ezek az adatsorok a következők voltak, a szimuláció forráskódjában is használt nevük szerint:

- **graph_pm**: Amikor egy újabb pont érkezik a gráfba, az minden alkalommal pontosan 1 ponthoz csatolódik hozzá. Ez az **array** típusú adatsor sorrendben tartalmazza, hogy az újabb csúcsok melyik másik csúcshoz kapcsolódtak hozzá beérkezésükkor. Hossza egyenlő az összes él darabszámával, mely jelen esetben egyel kevesebb a rendszerben szereplő pontok végleges számánál.
- **count_pm**: Ez a - szintén - **array** típusú adatsor tartalmazza, hogy az egyes pontok fokszámát. Az első eleme az első pont fokszáma, a második eleme a második pont fokszáma és így tovább. Hossza megegyezik az összes, rendszerben levő pont számával, ami pontosan egyel több, mint a rendszerben levő élek száma ebben az esetben.
- **dist_pm**: Ez az **array** tartalmazza sorrendben az egyes fokszámmal rendelkező csúcsok számát. Első eleme a 0 fokszámú csúcsok száma, második eleme az 1 fokszámú csúcsok száma, és így tovább. Hossza egyenlő a maximális lehetséges fokszámmal. Ez megegyezik az összes él számával, ugyanis ebben a kis valószínűségű esetben minden csúcs ugyanahhoz az egyetlen ponthoz csatlakozik. Ekkor a központi csúcshoz csatlakozik az összes E darab él, és így fokszáma is E .

A szimulációt három kezdőfeltétel esetére kellett vizsgálnunk, melyek a következők voltak:

- (1) Egy darab csúcs kezdetben
- (2) 5 darab csúcs, lineárisan összekötve. A második csúcs az elsőhöz, a harmadik a másodikhoz, és így tovább.
- (3) 5 darab csúcs, kereszt alakban összekötve. Négy csúcs az elsőhöz kötve, egymással semmilyen kapcsolatban nem állva.

2.2. Kiértékelés

Az (1)-es, (2)-es és (3)-as ábrákon a Lineáris Preferenciális Modell alapvető karakterisztikáit ábrázoltam, $E = 100$ él esetére, mindhárom kezdőfeltétel mellett. Nővekedő hálózat gyanánt ilyenkor $N = E + 1 = 101$ csúcs lesz a hálózatban, hisz minden alkalommal egy beérkező fa, pontosan 1 másik, már bent levő csúcshoz csatlakozik hozzá.

Az ábrákon a felső grafikon az egyes pontok végleges k fokszámait mutatja, a hálózatba történő beérkezési sorrendjükben, 1-től indexelve. A középső grafikonok az adott k fokszámú pontok N_k darabszámát mutatják, kezdve a $k = 0$ fokszámtól egészen az $E = 100$ esetében maximálisan lehetséges $k = 100$ -ig. Az alsó grafikonokon a $P_k = \frac{N_k}{N}$ diszkrét eloszlásfüggvényt ábrázoltam, mely tulajdonképpen csak a középső grafikon normáltja. Az utolsó függvény esetén vizsgálnunk kellett a $P_k = k^{-3}$ függvény, szimulált pontokra való illeszkedését. Ennek az eredményét a (4)-es grafikonon ábrázoltam. A grafikonon látható pontok a szimuláció mért értékei, a görbék pedig a rájuk illesztett köbös függvények. Az eltérés jelentéktelen volt sajnos. A pontosság javulását és az esetleges különbség feltárását az összes élszám növelésével lehetne elérni, azonban ehhez már túl nagy futásidőre lenne szükség.

A maximális foksám alakulását az élszám függvényében a (5)-ös ábrán illusztráltam. Az egyes görbék az egyes kezdőfeltételek mellett kapott értékeket jelölik az élszámok függvényében, $E \in [100, 2000]$ élszámértékekre. A vizsgált perióduson a grafikonon is láthatóan semmilyen eltérést nem észleltem a maximális értékekben, az esetleges különbségek is bőven a szóráshatáron belülre estek.

Az egyetlen különbség a grafikonok kezdeti értéke, azonban szinte azonnal együtt kezdenek haladni a továbbiakban. Tekintve a lassú növekedést és erős oszcillációt, eredetileg is hasonló eredményt vártam, hisz relatíve csak apró különbség volt a kezdeti feltételek között.

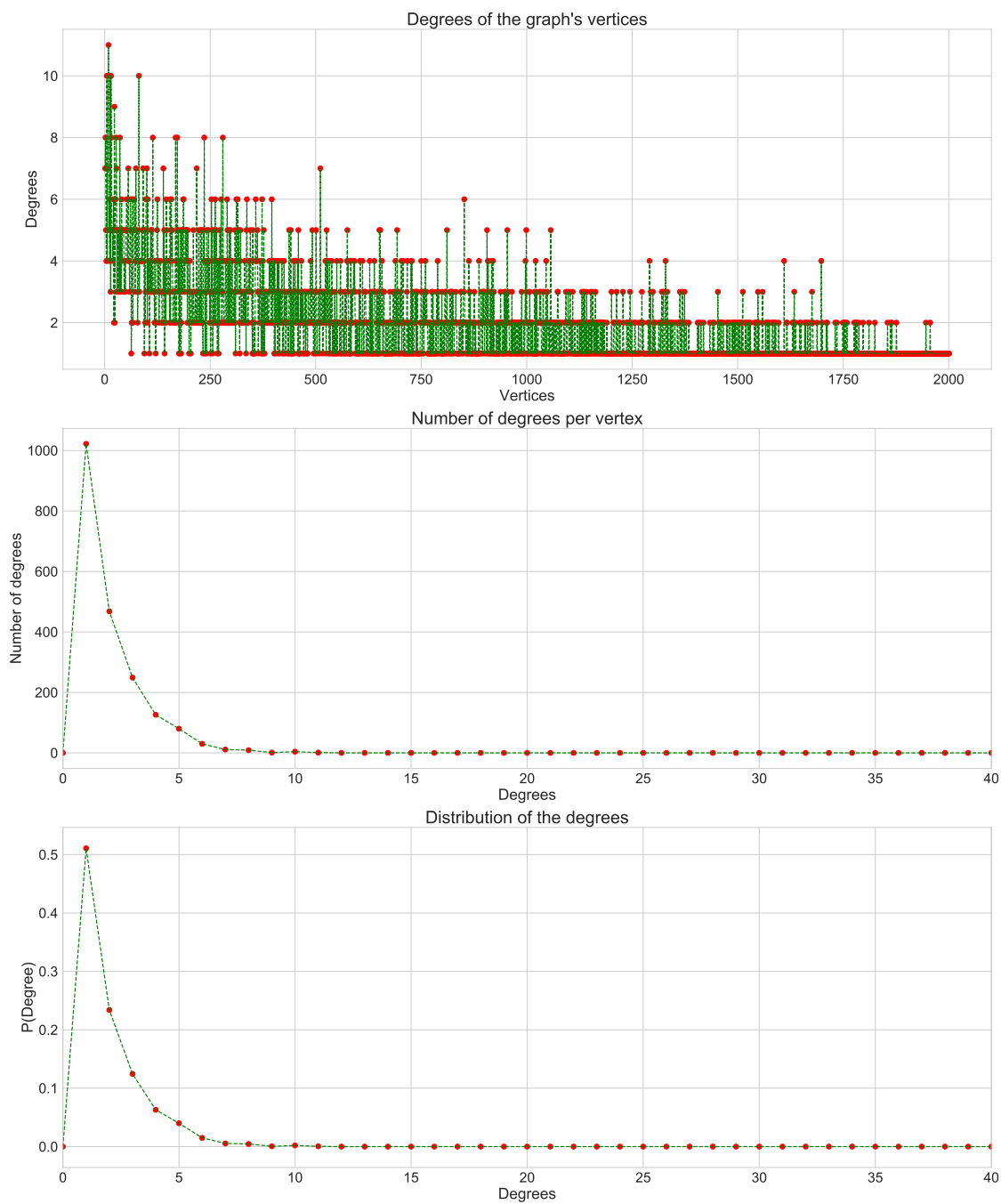
2.3. Diskuszió

A feladatok elvégzése után sajnos nem tudtam semmilyen különbséget kimutatni a különböző kezdőfeltételekkel történő futtatások között. Ha esetlegesen létezik valamilyen különbség ezen kezdőfeltételek között, akkor azt csakis az élszámok növelésével lehetne kimutatni. Különbség esetén egyes foksámokszámának aránya jobban megváltozna magasabb élszám esetén, így az eloszlás is megváltozna. Ezt sajnos nem sikerült egyértelműen kimutatni, vagy cáfolni.

Hivatkozások

- [1] Pál Balázs. *ELTE Computer Simulations 2019 — GitHub*. [Online; opened at May 1, 2019]. 2019. URL: https://github.com/masterdesky/ELTE_Stochastic_processes_2019.

APPENDIX A

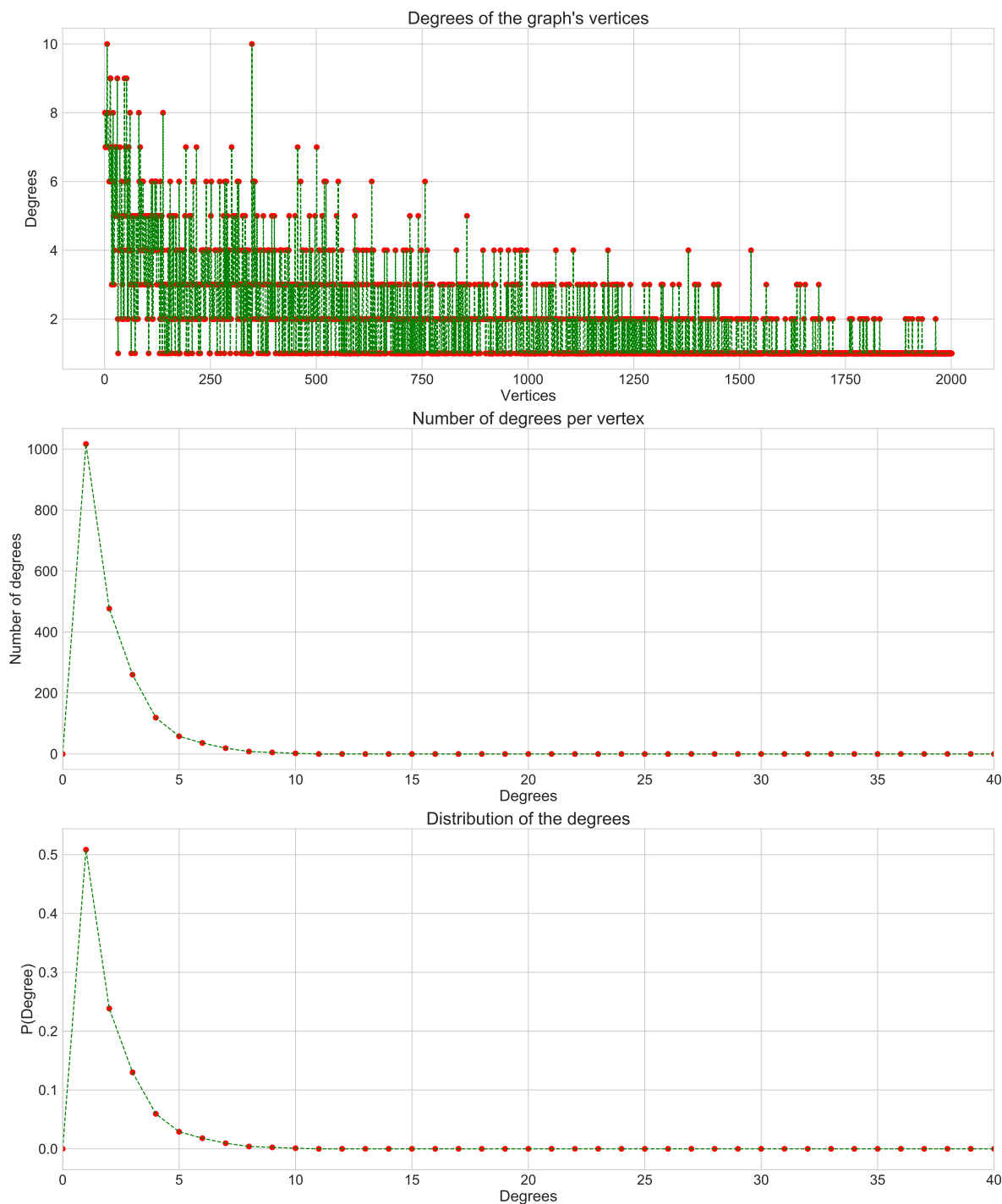


1. ábra. A Lineáris Preferenciális Modell $E = 100$ élre, 1. kezdőfeltétel

Fent: Az egyes csúcsok fokszáma

Középen: Adott fokszámú csúcsok darabszáma

Lent: A csúcsok fokszámának eloszlása

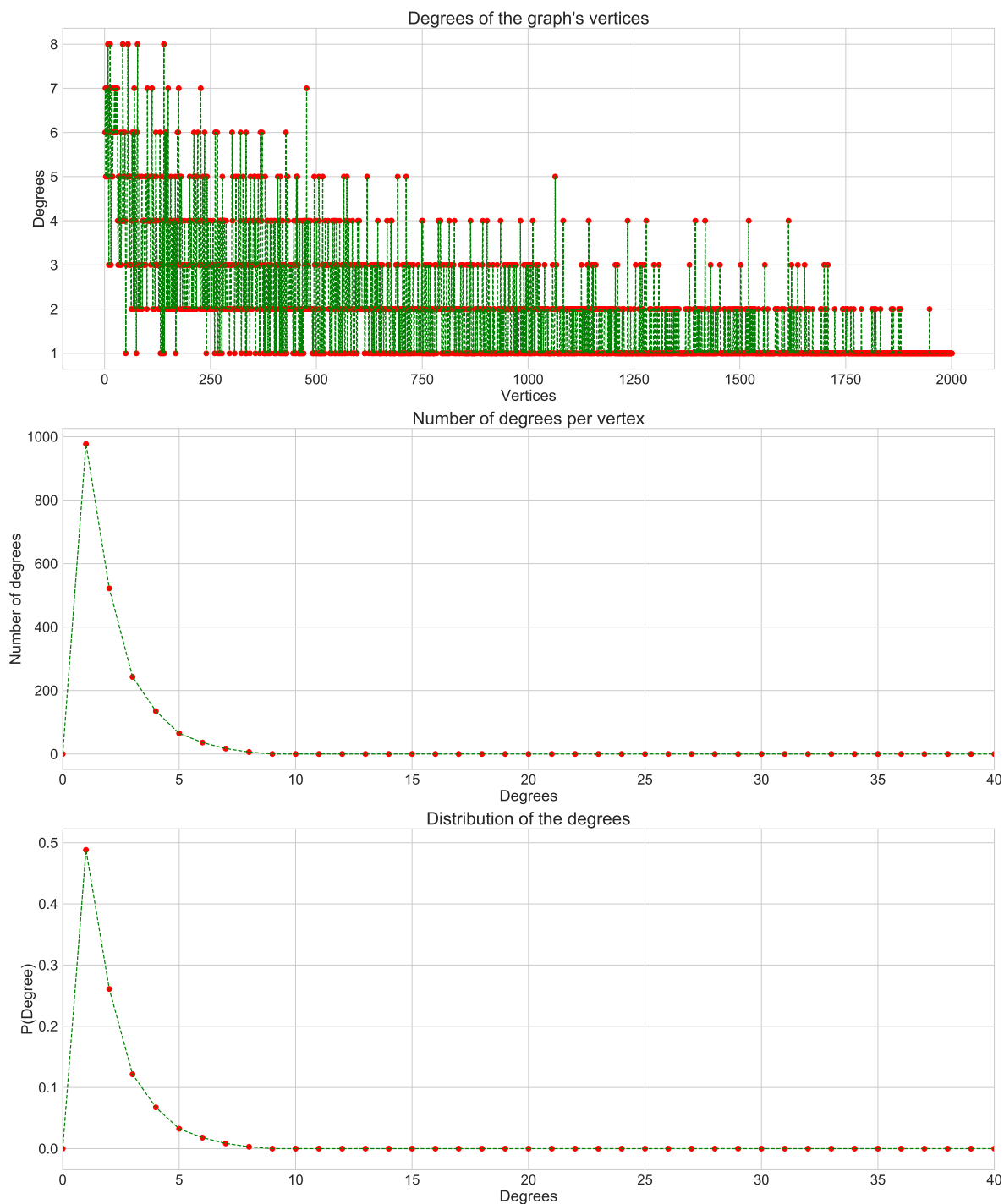


2. ábra. A Lineáris Preferenciális Modell $E = 100$ élre, 2. kezdőfeltétel

Fent: Az egyes csúcsok fokszáma

Középen: Adott fokszámú csúcsok darabszáma

Lent: A csúcsok fokszámának eloszlása

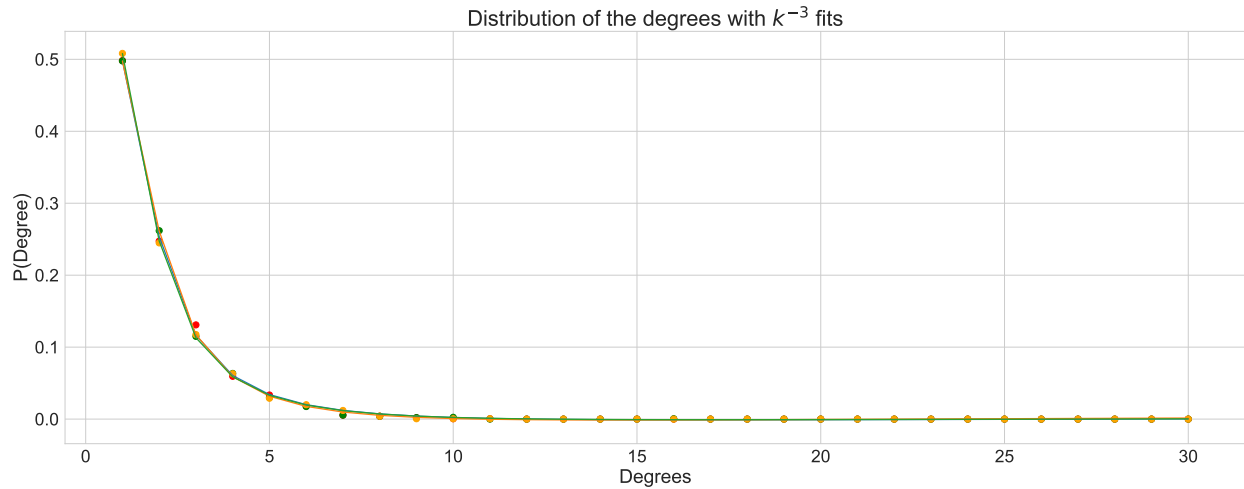


3. ábra. A Lineáris Preferenciális Modell $E = 100$ élre, 3. kezdőfeltétel

Fent: Az egyes csúcsok fokszáma

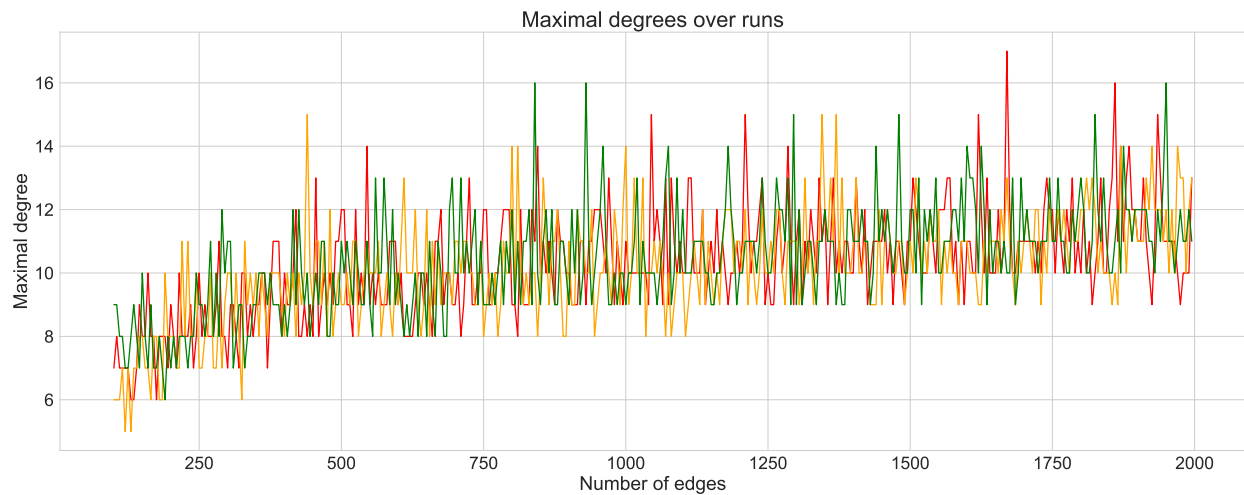
Középen: Adott fokszámú csúcsok darabszáma

Lent: A csúcsok fokszámának eloszlása



4. ábra. A Lineáris Preferenciális Modell különböző E élszámok esetére vett fokszámoszlása, mindhárom kezdőfeltétel mellett.

A képen látható pontok a szimulációban mért értékek, a görbék pedig a hozzájuk illesztett k^{-3} polinomok.



5. ábra. A Lineáris Preferenciális Modell különböző E élszámok esetén megjelenő maximális fokszáma.

A szimuláció $E = 100$ és $E = 2000$ élszámértékek között futott.

