Véletlen növekedő fák aszimptotikus vizsgálata TÉZISFÜZET

Rudas Anna

2012. december

Tartalomjegyzék

1. Bevezető

1.1. Modellcsalád és háttér

PhD disszertációmban egy véletlen növekedő fa modell aszimptotikus viselkedését vizsgálom, mely modell az úgynevezett *preferential attachment* tulajdonság egy általánosítáát valósítja meg.

Ebben a modellcsaládban a fa kezdetben pusztán egyetlen csúcsból, a gyökérből áll. Minden lépésben megjelenik egy-egy új csúcs, mely pontosan egy, már a fában jelen lévő csúcshoz csatlakozik. Azt a véletlen választást, hogy az új csúcs melyik másik csúcsot választja „szülőjéül”, a fában megfigyelhető fokszámok által meghatározott eloszlás szabályozza. Ezt a függést $w : \mathbb{N} \to \mathbb{R}_+$, az úgynevezett *súlyfüggvény* testesíti meg, mely paramétere a modellnek.

A modell definiálható diszkrét időben, ebben az esetben minden időlépésben egyetlen új csúcs születik, de definiálható folytonos időben is, akkor a csúcsok véletlen időpontokban születnek. Azon kérdések esetében, melyek vizsgálatunk tárgyát képezik, ez a két verzió ekvivalens módon átalakítható egymásba. A terület klasszikus modelljei és eredményei a diszkrét időben való megfogalmazást követik. Mi mégis a folytonos idejű megfogalmazással dolgozunk, mivel az általunk használni kívánt módszerek esetében ez a természetes és kézenfekvő.

Az úgynevezett *preferential attachment* modellek egyik legismertebb képviselője a Barabási–Albert féle gráf modell [3], melyben az új csúcs kapcsolódási hajlandósága pontosan arányos a már létező csúcsok fokszámaival. A Barabási–Albert modell fa esete, tehát amikor az új csúcs csak egyetlen már létező csúccsal kapcsolódik össze, a mi modellcsaládunknak speciális esete, nevezetesen w lineáris választásának felel meg. Valós hálózatokon megfigyelt tulajdonságokat, például a fokszám eloszlás hatványfüggvény szerinti lecsengését, sikeresen modellezi a Barabási–Albert féle modell. Ennek matematikailag precíz bizonyítását megtaláljuk Bollobás et al. [8]-ban, és tőlük függetlenül Móri [32] cikkében. Az ezekben a cikkekben alkalmazott technikák azonban olyan martingálok létezését használják, mely létezés csak a súlyfüggvény lineáris voltából fakadóan garantálható.

A *preferential attachment* fogalma általánosságban azt jelenti, hogy a w súlyfüggvény monoton növekedő. Az általunk vizsgált esetekben ez nem feltétlenül teljesül. Hasonló, általános súlyfüggvényeket korában már vizsgált Krapivsky és Redner [25] és [26], náluk $w(k) \sim k^\gamma$, és a precizitást ugyan mellőzve, de szerepelnek a két szétváló esetre, $\gamma > 1$ és $\gamma \leq 1$ vonatkozó eredmények. Az első esetben előbb-utóbb megjelenik egy domináns csúcs, mely azután majdnem minden új csúccsal összekapcsolódik, míg a többi csúcs fokszáma véges marad. Azok a súlyfüggvények, melyeket a jelen disszertációban mi tekintünk, nem okozzák a modell ilyen fajta „felrobbanását”, a mi modellcsaládunk a msódik, $\gamma \leq 1$ régióban marad.

Dereich és Mörters [11] munkájában a szerzők közelebbről megvizsgálják az egyes csúcsok fokszámának időbeli fejlődését, hasonló szublineáris tartományban, ahogyan mi. Ez a munka hivatkozza a mi [41] cikkünket. Másik hasonló, egyenletes és általánosított síkbeli rekurzív fa modelleket vizsgált korábban Smythe és Mahmoud [44].

A populáció növekedési modellek, melyeket az elágazó folyamatok elmélete részletesen vizsgál (lásd például Jagers [22]), szoros összefüggésben állnak a mi modellünkkel. Ez a kapcsolat szolgál sok, a jelen disszertációban szereplő bizonyítás alapjául.

A töredezési folyamatok által definiált fa növekedési folyamatok közeli összefüggést mutatnak a mi modellünk úgynevezett *globális* tulajdonságainak vizsgálatával. Ezekben a modellekben a fában fellelhető *távolságok* megfelelő átskálázása után a fa folyamat úgynevezett „véletlen valós fákhoz” illetve „folytonos véletlen fákhoz” konvergál. Ezen határobjektumok struktúráját részletesen feltárta pl. Haas, Miermont et al. [19, 20, 21].

A mi megközelítésünk, μ bevezetése a 4. fejezetben, különbözik ezektől. Ez egy mérték a végtelen, teljes fa levelein (melyben minden csúcsnak pontosan K gyereke van), ami ugyan metrikus tér, de triviális metrikával: nem térbeli skálázás eredménye, és nem is hordoz információt a fa növekedés folyamatáról. A mi esetünkben a μ által adott súlyok a fa méretének megfelelő átskálázásából adódnak, ahol méret alatt *számosságot* értünk. Összefoglalva, mi az aszimptotikus súlyeloszlást, nem pedig az aszimptotikus metrikus struktúrát vizsgáljuk. A súlyeloszlás határesetét a fizikus szakirodalomban is vizsgálták, lásd pl. Berestycki [5], itt a lokális dimenzióval analóg mennyiséget számolnak egy folytonos idejű töredezési folyamat esetére.

Hasonlóan, a Haas, Miermont et al. [19, 20, 21] által, a fejlődő fa térbeli skálázása után kapott folytonos fák esetében is, a metrikus struktúra képezi a vizsgálat tárgyát, és *halmazok* Hausdorff dimenziója, illetve Hausdorff mértéke a természetes kérdés, lásd Duquesne és Le Gall [15, 16]. Nálunk, ezzel ellentétben, nem a halmaz, hanem a *mérték* az, ami a fa fejlődésének hosszú távú leírását szolgálja, ennek megfelelően a mérték dimenzióját vizsgáljuk.

A fa növekedési folyamatunk folytonos idejű verziója elágazó bolyongássá is alakítható, amennyiben az idő szerepét az elmozdulás veszi át. Ekkor az aszimptotikus növekedés analóg módon vizsgálható, lásd Biggins tételét [7], vagy Lyons-nál [30]. Ezekben a munkákban ugyanakkor, mivel más a nézőpont, a határ struktúra esetében mások a természetesen felmerülő kérdések.

1.2. A disszertáció felépítése

A disszertáció három fő fejezetből épül fel, melyekben a modell definíció, illetve a tulajdonságok két családjának vizsgálata található. Jelen írás is ennek megfelelően bomlik három részre.

A 3. részben a fa *lokális* tulajdonságait vizsgáljuk: egy *tipikus*, azaz hosszú idő elteltével véletlenül választott csúcs környezetében nézünk körül. Ennek a szakasznak alapját két cikk képezi: [41], mely Valkó Benedekkel és Tóth Bálinttal, valamint [40], mely Tóth Bálinttal közös munka.

A 4. rész témája a modell *globális tulajdonságai*, itt a fa egészére tekintünk rá hosszú idő elteltével. Ez a szakasz [42]-ra épül, mely Tóth Imre Péterrel közös.

2. Jelölésrendszer, modell

2.1. Csúcsok, egyedek, fák

Gyökérrel rendelkező, rendezett fákat tekintünk, melyeket az irodalomban családfáknak vagy gyökérrel rendelkező síkbeli fáknak is neveznek.

A fákkal kapcsolatban kézenfekvő a családfák esetében megszokott szóhasználat. A fára tehát egy populáció fejlődésének kódolásaként tekintünk, mely egyetlen egyedből, a gyökérből ered, akinek a „gyerekeiből” áll az első generáció, ezek a csúcsok azok, melyek közvetlenül, éllel csatlakoznak a gyökérhez. Általában véve, az élek tehát apa-fiú kapcsolatot reprezentálnak, amiben a szülő mindig az, aki a gyökérhez közelebb található. A testvérek közötti születési sorrendet is figyelembe vesszük, ez nyilvánul meg abban, hogy a fa rendezett (síkbeli).

Rögzítsük a pozitív egész számok egy $\mathbb{I}$ részhalmazát, és címkézzük a gyökeres, rendezett fa csúcsait a következő halmaz elemeivel,

$$\mathcal{N} := \bigcup_{n=0}^{\infty} \mathbb{I}^n, \quad \text{ahol} \quad \mathbb{I}^0 := \{\emptyset\}\,. \tag{1}$$

Kissé eltérő modell eseteket tárgyalunk a 3. és 4. szakaszokban, emiatt ezekben $\mathbb{I}$ definíciója különbözőképp alakul majd, a következőképpen.

- A 3. szakaszban a $\mathbb{I} = \mathbb{Z}^+$ választással élünk, ami azt jelenti, hogy bármelyik csúcs tetszőleges számú gyerekkel rendelkezhet.
- A 4. szakaszban rögzítünk egy $K \in \mathbb{N}, ; K \geq 2$ pozitív egész számot, és a $\mathbb{I} := = \{1,2,\ldots,K\}$ választjuk. Ez tehát azt eredményezi, hogy minden csúcsnak legfeljebb K gyereke lehet.

Jelölésrendszerünk szerint $\emptyset$ jelenti a gyökeret, a gyermekei, születésük sorrendjében $\mathbb{I}$ elemeivel jelöltek, és általában $x = (x_1, x_2, \ldots, x_k) \in \mathcal{N}$ gyermekeit $(x_1, x_2, \ldots, x_k, 1)$, $(x_1, x_2, \ldots, x_k, 2), \ldots$ jelölik. Tehát ha egy csúcs címkéje $x = (x_1, x_2, \ldots, x_k) \in \mathcal{N}$, akkor ez azt jelenti, hogy ő a szülőjének a x_k-adik gyereke, aki pedig a saját szülőjének a $x_{k-1}-$edik gyereke, és így tovább.

A gyökeres, rendezett fákat a csúcsainak címkéivel azonosítjuk, hiszen ez már minden információt tartalmaz az élekről is. Világos, hogy egy $G \subset \mathcal{N}$ részhalmaz pontosan akkor jelent egy gyökeres, rendezett fát, hogyha $\emptyset \in G$, továbbá minden $(x_1, x_2, \ldots, x_k) \in G$ esetében $(x_1, x_2, \ldots, x_{k-1}) \in G$ is teljesül, valamint $(x_1, x_2, \ldots, x_k - 1) \in G$, ha $x_k > 1$.

Az összes véges, gyökeres, rendezett fák halmazát jelöljük $\mathcal{G}$-vel. Úgy tekintünk egy $G \in \mathcal{G}$-re, mint irányított fára, melynek élei a szülők felől a gyerekek felé mutatnak. Egy $x \in G$ csúcs *foka* a gyermekeinek számát jelenti G-ben, tehát ez a terminológia kissé különbözik a szokásostól:

$$\deg(x, G) := \max\{n \in \mathbb{I} : xn \in G\}.$$

$G \in \mathcal{G}$ n-edik *generációja*

$$G_{[n]} := \{x \in G : |x| = n\}, \quad n \geq 0,$$

ahol $|x| = n$ pontosan akkor, ha $x \in \mathbb{I}^n$.

Legyen $k \geq n$, ekkor egy $x = (x_1, x_2, \ldots, x_k) \in \mathcal{N}$ csúcs *n-edfokú őse* $x^n = (x_1, x_2, \ldots, x_{k-n})$ ha $k > n$, és $x^n = \emptyset$, ha $k = n$.

Egy $x \in G$ csúcs alatti *részfa*:

$$G_{\downarrow x} := \{y : xy \in G\}, \tag{2}$$

ez tehát az x utódaiból álló részfa, mint x gyökerű, rendezett fa. Ha az $x = (x_1, x_2, \ldots, x_n) \in$ $\in \mathcal{N}$ csúcsra $|x| = n \geq k$, akkor használjuk a $x_{\downarrow k} = (x_{n-k+1}, x_{n-k+2}, \ldots, x_n)$ jelölést. Ez az a címke, amit az $x \in G$ csúcs kapna a $G_{\downarrow x^k}$ részfában.

2.2. A véletlen fa modell

A véletlen fa modell paramétereként a $w : \mathbb{N} \to \mathbb{R}_+$ súlyfüggvényt rögzítjük.

A diszkrét idejű modell definiálásához nincs szükségünk további, w-re vonatkozó megszorításokra. A folytonos idejű modell esetében azonban rászorulunk, hogy az (M) feltételezéssel éljünk, lásd később ebben a szakaszban. Ez a modell definíciójához éppúgy kell, mint az 3. szakaszban tárgyalt eredményeinkhez. A 4. szakaszban megköveteljük, hogy $w(k) = 0, \quad k \geq K$, ami egyfelől biztosítja, hogy minden csúcsnak legfeljebb K gyereke lesz, másfelől azt is, hogy automatikusan teljesül az (M) feltétel.

Diszkrét idő

Adott $w : \mathbb{N} \to \mathbb{R}_+$ súlyfüggvény mellett definiáljuk a következő, diszkrét idejű Υ^d Markov-láncot, a $\mathcal{G}$ megszámlálható állapottéren, $\Upsilon^d(0) = \{\emptyset\}$ kezdeti állapottal. Ha valamely $n \geq 0$-re $\Upsilon^d(n) = G$, akkor az $x \in G$ csúcsra legyen $k := \deg(x, G) + 1$. Ezt a jelölésrendszert használva az átmenetvalószínűségek legyenek

$$\mathbf{P}(\Upsilon^d(n+1) = G \cup \{xk\}) = \frac{w(\deg(x, G))}{\sum_{y \in G} w(\deg(y, G))}.$$

Más szavakkal, minden diszkrét időlépésben egy új csúcs érkezik, és hozzácsatlakozik pontosan egy, már a fában jelen lévő csúcshoz. Hogyha az adott időpontban a véletlen fa G-vel egyenlő, akkor az $x \in G$ csúcs választásának valószínűsége $w(\deg(x, G))$-vel arányos.

Folytonos idő

Adott $w : \mathbb{N} \to \mathbb{R}_+$ súlyfüggvény mellett legyen $X(t)$ egy Markovi, tiszta születési folyamat $X(0) = 0$ kezdeti állapottal és a következő születési rátákkal,

$$\mathbf{P}\big(X(t + \mathrm{d}t) = k + 1 \,\big|\, X(t) = k\big) = w(k)\,\mathrm{d}t + o(\mathrm{d}t).$$

Legyen $\rho : [0,\infty) \mapsto (0,\infty]$ a $X(t)$ születési folyamathoz tartozó sűrűség, azaz

$$\rho(t) = \lim_{\varepsilon\to 0} \varepsilon^{-1}\mathbf{P}((t, t+\varepsilon) \text{ tartalmaz pontot } X\text{-ből}) .$$

Legyen $\widehat{\rho} : (0,\infty) \to (0,\infty]$ a ρ sűrűség (formális) Laplace transzformáltja:

$$\widehat{\rho}(\lambda) := \int_0^\infty e^{-\lambda t}\rho(t)\,\mathrm{d}t = \sum_{n=0}^{\infty}\prod_{i=0}^{n-1}\frac{w(i)}{\lambda + w(i)}.$$

Legyen továbbá

$$\underline{\lambda} := \inf\{\lambda > 0 : \widehat{\rho}(\lambda) < \infty\}.$$

A disszertáció teljes egészében megköveteljük, hogy a w súlyfüggvény eleget tegyen a következő feltételnek:

$$\lim_{\lambda\searrow\underline{\lambda}} \widehat{\rho}(\lambda) > 1. \tag{M}$$

Most már készen állunk arra, hogy definiáljuk a $\Upsilon(t)$ folyamatot, ami tehát egy folytonos idejű, időben homogén átmenetű Markov-folyamat a $\mathcal{G}$ megszámlálható állapottéren, a $\Upsilon(0) = \{\emptyset\}$ kezdeti állapotból indítva.

Az ugrási ráták a következők: ha egy $t \geq 0$ időpontban $\Upsilon(t) = G$, akkor a folyamat a $G \cup \{xk\}$ állapotba ugorhat, $w(\deg(x,G))$ rátával, ahol $x \in G$ és $k = \deg(x,G)+1$. Ez másképp fogalmazva azt jelenti, hogy minden már létező $x \in \Upsilon(t)$ csúcs a többiektől függetlenül, $w(\deg(x,\Upsilon(t)))$ rátával 'életet ad egy gyereknek'.

Vegyük észre, hogy az (M) feltétel biztosítja, hogy

$$\sum_{k=0}^{\infty}\frac{1}{w(k)} = \infty,$$

így a $\Upsilon(t)$ Markov-lánc jóldefiniált $t \in [0,\infty)$-re, azaz nem robban fel véges idő alatt.

Az x csúcs születésének időpontját jelöljük τ_x-szel:

$$\tau_x := \inf\{t > 0 \ : \ x \in \Upsilon(t)\} .$$

Kapcsolat a folytonos és diszkrét idő között

Ha csak azokban a megállási időkben tekintünk a folyamatra, amikor egy-egy új csúcs születik,

$$T_n := \inf\{t : |\Upsilon(t)| = n+1\},$$

akkor épp a diszkrét idejű modellt kapjuk: $\Upsilon(T_n)$ ugyanolyan eloszlású, mint $\Upsilon^d(n)$.

3. Lokális tulajdonságok

3.1. Kérdésfeltevés, háttér

Ebben a szakaszban a véletlen fa *lokális tulajdonságait* vizsgáljuk, hosszú idővel a folyamat kezdete után. A „tipikus", azaz egyenletesen véletlenül választott csúcs környezetével kapcsolatban teszünk fel kérdéseket.

Főbb eredményeink a következők. Megadjuk a fokszámok aszimptotikus eloszlását, azaz egy (egyenletesen) véletlenül választott csúcs fokszámának határeloszlását. Mélyebben körülnézve a véletlen csúcs környezetében, meghatározzuk ezen kívül a véletlen csúcs alatt található részfa határeloszlását is. Végezetül megadjuk a teljes fa aszimptotikus eloszlását is, a véletlen csúcsból nézve.

A módszerünk kulcsa, hogy a modellt folytonos időbe ágyazzuk be, ahogy azt a 2.2 szakaszban már meg is tettük. A legfőbb előnye ennek a beágyazásnak kétségkívül az, hogy rávilágít az eredeti, diszkrét idejű fa modell és az elágazó folyamatok jól kidolgozott elméletének kapcsolatára. A legfőbb *lokális* eredményeink ez által a kapcsolat által nyernek bizonyítást. Hasonló kapcsolat feltárása révén jutott el eredményeihez korábban B. Pittel [38], ebben a szerző egy Crump–Mode elágazó folyamat segítségével bizonyítja egyenletes és általánosított síkbeli fák, és véletlen m-adikus keresőfák magasságáról szóló tételeit.

3.2. Tézisek

Az (M) feltétel következményeképp a

$$\widehat{\rho}(\lambda) = 1 \tag{3}$$

egyenletnek létezik egyetlen λ^* megoldása.

Most már kimondhatjuk az első tételünket.

3.1. Theorem. *Legyen w az* (M) *feltételnek eleget tevő súlyfüggvény, és legyen λ^* mint fent. Vegyünk egy tetszőleges, korlátos $\varphi : \mathcal{G} \to \mathbb{R}$ tesztfüggvényt. Ekkor a következő határérték eléretik majdnem biztosan:*

$$\lim_{t\to\infty} \frac{1}{|\Upsilon(t)|} \sum_{x\in\Upsilon(t)} \varphi(\Upsilon(t)_{\downarrow x}) = \lambda^* \int_0^\infty e^{-\lambda^* t}\mathbf{E}\big(\varphi(\Upsilon(t))\big)\,\mathrm{d}t.$$

A 3.1 tételből a fa aszimptotikus tulajdonságairól számos eredmény következik , *a fát egy véletlenül választott ζ csúcsból nézve*, amit egyenletesen választunk $\Upsilon(t)$-ből. Megadjuk a véletlenül választott csúcs fokszámának határeloszlását, a véletlen csúcs alatt található részfa határeloszlását, valamint a véletlen csúcs k-adik őse alatti részfa eloszlását is. Képlettel: meghatározzuk $\deg(\zeta, \Upsilon(t)) \in \mathbb{N}$, $\Upsilon(t)_{\downarrow\zeta} \in \mathcal{G}$ és $(\Upsilon(t)_{\downarrow\zeta^{(k)}}, \zeta_{\downarrow k}) \in \mathcal{G}^{(k)}$ határeloszlását.

Annak érdekében, hogy pontosan megfogalmazhassuk a 3.1. Tétel ezen következményeit, bevezetünk még néhány jelölést.

A $\mathcal{G}$ halmazon értelmezett $\boldsymbol{\pi}$ valószínűségi mértéket *kiegyensúlyozottnak* mondunk, ha

$$\sum_{H\in\mathcal{G}}\boldsymbol{\pi}(H)\sum_{x\in H_{[1]}}\mathbb{1}\{H_{\downarrow x}=G\}=\boldsymbol{\pi}(G). \tag{4}$$

Legyen ezen kívül rögzítve $G\in\mathcal{G}$ és annak valamelyik történelmi sorrendje, $s=(s_0,s_1,\ldots,s_{|G|-1}) \in\mathcal{S}(G)$. Az ehhez a történelemhez tartozó totális súlyt definiáljuk mint

$$W(G,s,i):=W(G(s,i))$$

minden $0\le i\le|G|-1$ egészre, és a közben megjelenő csúcsok megfelelő súlyai legyenek

$$w(G,s,i):=w\left(\deg\left((s_i)^1,G(s,i-1)\right)\right).$$

Mivel $\deg\left((s_i)^1,G(s,i-1)\right)$ s_i szülőjének foka, épp mielőtt s_i megjelent volna, így $w(G,s,i)$ az a ráta, amivel a véetlen fa folyamat $G(s,i-1)$-ből $G(s,i)$-be ugrik.

Legyen w az (M) feltételnek eleget tevő súlyfüggvény, és legyen λ^* mint fent. Ezek mellett definiáljuk a következőket,

$$\begin{aligned}\boldsymbol{p}_w(k) &:= \frac{\lambda^*}{\lambda^*+w(k)}\prod_{i=0}^{k-1}\frac{w(i)}{\lambda^*+w(i)},\\ \boldsymbol{\pi}_w(G) &:= \sum_{s\in\mathcal{S}(G)}\frac{\lambda^*}{\lambda^*+W(G)}\prod_{i=0}^{|G|-2}\frac{w(G,s,i+1)}{\lambda^*+W(G,s,i)}.\end{aligned}$$

3.2. Theorem. *Legyen w az* (M) *feltételnek eleget tevő súlyfüggvény, és legyen λ^* mint fent. Ekkor a következő határértékek eléretnek majdnem biztosan:*

(a) Bármely rögzített $k\in\mathbb{N}$-re

$$\lim_{t\to\infty}\frac{|\{x\in\Upsilon(t):\deg(x,\Upsilon(t))=k\}|}{|\Upsilon(t)|}=\boldsymbol{p}_w(k).$$

(b) Bármely rögzített $G\in\mathcal{G}$-re

$$\lim_{t\to\infty}\frac{\left|\{x\in\Upsilon(t):\Upsilon(t)_{\downarrow x}=G\}\right|}{|\Upsilon(t)|}=\boldsymbol{\pi}_w(G).$$

(c) Bármely rögzített $(G,u)\in\mathcal{G}^{(k)}$-ra

$$\lim_{t\to\infty}\frac{\left|\{x\in\Upsilon(t):(\Upsilon(t)_{\downarrow x^{(k)}},x_{\downarrow k})=(G,u)\}\right|}{|\Upsilon(t)|}=\boldsymbol{\pi}_w(G).$$

Mindezeken felül $\boldsymbol{p}_w$ és $\boldsymbol{\pi}_w$ valószínűségi eloszlások az $\mathbb{N}$ és $\mathcal{G}$ halmazokon, és $\boldsymbol{\pi}_w$ kiegyensúlyozott (azaz érvényes a (4) *feltétel).*

4. Globális tulajdonságok

4.1. Kérdésfeltevés, háttér

A következő kérdés természetes módon merül fel. Rögzítsünk egy csúcsot, például az első gyereket az első generációban (a gyökér első gyerekét). Mi ennek a csúcsnak az „aszimptotikus népszerűsége", a generációjában megjelenő többi csúcshoz viszonyítva? Azt értjük ezalatt, hogy hány leszármazottja lesz ennek a csúcsnak, hosszú idő múlva, a testvérei leszármazottjainak számához viszonyítva.

Ugyanennek a kérdésnek másik oldalról való megfogalmazása, hogy rögzítünk egy csúcsot, várunk hosszú ideig, majd választunk a véletlen fából egyenletesen egy csúcsot. A kérdés, mi a valószínűsége, hogy az ily módon választott véletlen csúcs a korábban rögzített csúcs leszármazottja lesz? Világos, hogy ha ezeket a határ valószínűségeket tekintjük például az első generációban, akkor egy olyan eloszlást kapunk, ami maga is véletlen, és a fa növekedéséről árul el valamit.

Ha tekintjük ezen véletlen határ mértékek rendszerét (amint az idő a végtelenbe tart, de a generációk véges, fix értékek), érdekesnek tűnik feltenni a kérdést, mondhatunk-e valamit, hogyha a generációs szintet is elengedjük a végtelenbe. Ezt az ötletet precízzé téve fogjuk bevezetni a μ határmértéket.

Amikor megalkottuk a véletlen határmértéket, ami a végtelen rendszer aszimptotikus viselkedéséről tanúskodik, természetes, hogy megvizsgáljuk ennek a mértéknek a *Hausdorff (és pakolási) dimenzióját*. Másrészt, a mérték dimenziója a metrikán is múlik, amiben viszont van egy tetszőleges konstans. Hogy kiszűrjük ezt a triviális függést, általánosan használt megközelítés, hogy a *határmérték entrópiájáról* folytassuk a vizsgálatot, ami már csak magától a növekedési folyamattól függ. Ez tehát a dimenzióval analóg fogalom a növekedés szempontjából.

Az eredményünk kulcsa az a Markov lánc, ami egy μ-tipikus levél megkonstruálása folyamán természetes módon megjelenik. A fa szerkezetének vizsgálata során a Markov tulajdonság hamar előbukkan, némi technikai nehézség abból fakad, hogy az állapottér nem kompakt.

A modell választás speciális abból a szempontból, hogy csak véges fokszámot engedünk meg a csúcsokon, de abban a tekintetben általános, hogy K rögzítése után az w súlyfüggvény bármilyen, a $\{0,1,\ldots,K-1\}$ halmazon értelmezett pozitív értékű függvény lehet.

4.2. További jelölések, és w megválasztása

A modellünk paraméteréül szolgáló súyfüggvénytől megköveteljük az alábbiakat.

Rögzítünk egy $K>2$ egész számot, ami fölött w nulla lesz:

$$w(k)=0, \quad k \geq K. \tag{5}$$

Ez a megszorítás azt eredményezi, hogy minden csúcsnak legfeljebb K gyereke lehet. A véletlen fa csúcsai ilyetén módon a következő halmaz elemeivel címkézettek,

$$\mathcal{N} := \bigcup_{n=0}^{\infty} \mathbb{I}^n,$$

ugyanúgy mint korábban (1)-ben, csak éppen most $\mathbb{I}=\{1,2,\ldots,K\}$. Mivel megköveteljük (5) teljesülését, ezért az (M) feltétel is automatikusan teljesül.

4.3. Határ objektumok

Mint azt (3)-ban láttuk, a

$$\widehat{\rho}(\lambda) = 1$$

egyenletnek létezik egyetlen $\lambda^* > 0$ megoldása, a Malthus-i paraméter. Ez a λ^* adja meg majdnem biztosan a fa méretének exponensét: a méret normalizáltja majdnem biztosan tart egy valószínűségi változóhoz, amit Θ jelöl:

$$\Theta := \lim_{t\to\infty} e^{-\lambda^* t} |\Upsilon(t)| \, .$$

Minden $x \in \mathcal{N}$ csúcshoz vezessük be az analóg változót Θ_x néven,

$$\Theta_x := \lim_{t\to\infty} e^{-\lambda^*(t-\tau_x)} \, |\Upsilon_{\downarrow x}(t)| \; .$$

Világos, hogy minden $x \in \mathcal{N}$ csúcsra a Θ_x valószínűségi változók azonos eloszlásúak. Az összefüggés ezek között az, hogy bármely $x \in \mathcal{N}$-re

$$\Theta_x = \sum_{i=1}^{K} e^{-\lambda^*(\tau_{xi}-\tau_x)} \Theta_{xi} \, ,$$

ami egyenesen következik abból, hogy $|\Upsilon_{\downarrow x}(t)| = 1 + \sum_{i=1}^{K} |\Upsilon_{\downarrow xi}(t)|$.

Most tegyük fel a következő kérdést. Rögzítsük $x \in \mathcal{N}$ csúcsot, és a t időpontban válasszunk egy ζ_t csúcsot egyenletesen véletlenül $\Upsilon(t)$-ből. Mi a valószínűsége annak, hogy ζ_t az x csúcs leszármazottja? Ahogy azt (6)-ban alább láthatjuk, ez a valószínűség majdnem biztosan tart a Δ_x határértékhez, amint $t \to \infty$, és Δ_x kifejezhető a τ and Θ valószínűségi változók segítségével,

$$\Delta_x := \lim_{t\to\infty} \frac{|\Upsilon_{\downarrow x}(t)|}{|\Upsilon(t)|} = e^{-\lambda^*\tau_x} \lim_{t\to\infty} \frac{e^{-\lambda^*(t-\tau_x)}|\Upsilon_{\downarrow x}(t)|}{e^{-\lambda^* t}|\Upsilon(t)|} = \frac{e^{-\lambda^*\tau_x}\Theta_x}{\Theta_\emptyset} \, . \tag{6}$$

Most bármely $n \in \mathbb{N}$ esetén definiálhatjuk a $\mathbb{I}^n = \{x : |x| = n\}$ generáción értelmezett μ_n véletlen mértéket,

$$\mu_n(\{x\}) := \Delta_x \, .$$

Ez majdnem biztosan egy valószínűségi mérték, ami abból következik, hogy $\Delta_\emptyset = 1$ és $\Delta_y = \sum_{i=1}^{K} \Delta_{yi}$.

Jelölje H_n a μ_n mérték entrópiáját, azaz

$$H_n = -\sum_{|x|=n} \Delta_x \log \Delta_x \, .$$

Véletlen mérték, mint a fa határ objektuma

Jelölje $\partial\mathcal{N}$ a teljes, végtelen fa leveleit: $\partial\mathcal{N} = \{1,2,\ldots,K\}^\infty$. Adott $x \in \mathcal{N}$ és $y \in \partial\mathcal{N}$ csúcsok esetén az xy konkatenáció értelmes, és az eredménye $xy \in \partial\mathcal{N}$. Valamely $x \in \mathcal{N}$ és $z \in \partial\mathcal{N}$ esetén $x \prec z$ jelölje, ha $\exists y \in \partial\mathcal{N}$ hogy $z = xy$. Adott $x \in \mathcal{N}$ csúcs esetén jelölje az x alatti leveleket $\partial\mathcal{N}(x) = \{z \in \partial\mathcal{N} : x \prec z\}$. Jelölje $x_{|l}$ az x szó első l betűjét, vagy más szavakkal, jelölje x azon ősét, aki a fa l-edik generációjában van, és vezessük be $\partial\mathcal{N}$-en a szokásos metrikát,

$$d(x,y) = \Lambda^{\max\{n\in\mathbb{N} : x_{|n}=y_{|n}\}}\,, \tag{7}$$

valamilyen $0<\Lambda<1$ kontrakciós konstanssal. Ezt a konstanst gyakran $1/e$-nek választják, ami egyszerűsít bizonyos képleteket. Mi nem rögzítjük Λ értékét, annak érdekében, hogy világosan be tudjuk majd azonosítani a leendő képleteinkben az ettől való függést.

A μ_n véletlen mértékek segítségével defniáljuk μ-t $\partial\mathcal{N}$-nek $\partial\mathcal{N}(x)$ cilinderhalmazain,

$$\mu(\partial\mathcal{N}(x)) := \mu_n(\{x\}) = \Delta_x\ , \text{ if } |x| = n\ ,$$

majd pedig a szokásos módon kiterjesztjük μ-t $\{\partial\mathcal{N}(x) : x \in \mathcal{N}\}$-ről a generált szigma-algebrára. A tételeink erre a kiterjesztett μ mértékre vonatkoznak.

4.4. Tézisek

4.1. Theorem. *A határ entrópia*

$$h := \lim_{n\to\infty} \frac{1}{n} H_n$$

létezik és egy valószínűséggel konstans.

4.2. Theorem. *A μ mérték $\dim_H \mu$ Hausdorff dimenziója és $\dim_P \mu$ pakolási dimenziója egy valószínűséggel megegyezik és konstans, valamint h-ra és a dimenzióra teljesül, hogy*

$$\dim_H \mu = \dim_P \mu = \frac{h}{-\log\Lambda},$$

ahol Λ a (7)-ben definiált kontrakció. Ezen felül μ lokális dimenziója μ-majdnem minden pontban megegyezik a $\dim_H \mu = \dim_P \mu$ értékkel.

4.3. Theorem. *Explicit formula adható h értékére:*

$$h = \mathbf{E}\left(\sum_{i=1}^{K} \lambda^* \tau_i e^{-\lambda^* \tau_i}\right).$$

Ez az érték a w súlyfüggvény ismeretében számolható.

Hivatkozások

[1] David Aldous. *Probability Approximations via the Poisson Clumping Heuristic*. Springer, 1989.

[2] David Aldous. Asymptotic fringe distributions for general families of random trees. *Ann. Appl. Probab.*, 1(2):228–266, 1991.

[3] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[4] Itai Benjamini and Oded Schramm. Recurrence of distributional limits of finite planar graphs. *Electronic J. Probab.*, 6(paper 23):1–13, 2001.

[5] Julien Berestycki. Multifractal spectra of fragmentation processes. *Journal of Statistical Physics*, 113(3):411–430, 2003.

[6] Jean Bertoin. *Random fragmentation and coagulation processes*. Cambridge Univ Pr, 2006.

[7] J. D. Biggins. Martingale convergence in the branching random walk. *Journal of Applied Probability*, 14(1):25–37, 1977.

[8] Béla Bollobás, Oliver Riordan, Joel Spencer, and Gábor Tusnády. The degree sequence of a scale-free random graph process. *Random Structures Algorithms*, 18(3):279–290, 2001.

[9] Béla Bollobás and Oliver M. Riordan. Mathematical results on scale-free random graphs. In *Handbook of graphs and networks*, pages 1–34. Wiley-VCH, Weinheim, 2003.

[10] Fan Chung, Shirin Handjani, and Doug Jungreis. Generalizations of Polya's urn problem. *Ann. Comb.*, 7(2):141–153, 2003.

[11] Steffen Dereich and Peter Mörters. Random networks with sublinear preferential attachment: degree evolutions. *Electron. J. Probab.*, 14(43):1222–1267, 2009.

[12] Rui Dong, Christina Goldschmidt, and James B. Martin. Coagulation-fragmentation duality, Poisson-Dirichlet distributions and random recursive trees. *Ann. Appl. Probab.*, 16(4):1733–1750, 2006.

[13] Joseph Leo Doob. *Stochastic Processes*. Wiley, 1953.

[14] Michael Drmota. *Random trees*. SpringerWienNewYork, Vienna, 2009. An interplay between combinatorics and probability.

[15] T. Duquesne. Packing and Hausdorff measures of stable trees. *Lévy Matters I*, pages 93–136, 2010.

[16] Thomas Duquesne and Jean-FranĂ§ois Le Gall. Probabilistic and fractal aspects of Lévy trees. *Probability Theory and Related Fields*, 131:553–603, 2005.

[17] Richard Durrett. *Random Graph Dynamics.* Cambridge University Press, Cambridge, 2007. Cambridge Series in Statistical and Probabilistic Mathematics.

[18] Kenneth Falconer. *Techniques in Fractal Geometry.* Wiley, 1997.

[19] B. Haas and G. Miermont. The genealogy of self-similar fragmentations with negative index as a continuum random tree. *Electronic Journal of Probability*, 9(paper 4):57, 2004.

[20] B. Haas and G. Miermont. Scaling limits of Markov branching trees, with applications to Galton-Watson and random unordered trees. 2010. Arxiv preprint, `http://arxiv.org/abs/1003.3632`.

[21] B. Haas, G. Miermont, J. Pitman, and M. Winkel. Continuum tree asymptotics of discrete fragmentations and applications to phylogenetic models. *The Annals of Probability*, 36(5):1790–1837, 2008.

[22] Peter Jagers. *Branching processes with biological applications.* Wiley-Interscience [John Wiley & Sons], London, 1975. Wiley Series in Probability and Mathematical Statistics —Applied Probability and Statistics.

[23] Peter Jagers and Olle Nerman. The growth and composition of branching populations. *Adv. in Appl. Probab.*, 16(2):221–259, 1984.

[24] H. Kesten and B. P. Stigum. A limit theorem for multidimensional Galton-Watson processes. *Ann. Math. Statist.*, 37:1211–1223, 1966.

[25] P. L. Krapivsky and S. Redner. Organization of growing random networks. *Phys. Rev. E*, 63(6):066123, May 2001.

[26] P. L. Krapivsky, S. Redner, and F. Leyvraz. Connectivity of growing random networks. *Phys. Rev. Lett.*, 85(21):4629–4632, Nov 2000.

[27] Norbert Kusolitsch. Why the theorem of Scheffé should be rather called a theorem of Riesz. *Period. Math. Hungar.*, 61(1-2):225–229, 2010.

[28] Jean-François Le Gall. Processus de branchement, arbres et superprocessus. In *Development of mathematics 1950–2000*, pages 763–793. Birkhäuser, Basel, 2000.

[29] László Lovász. *Very large graphs.* 2008. Arxiv preprint, `http://arxiv.org/abs/0902.0132`.

[30] R. Lyons. A simple path to Biggins' martingale convergence for branching random walk. In K.B. Athreya and P. Jagers, editors, *Classical and modern branching processes*, The IMA volumes in mathematics and its applications. Springer, 1997.

[31] R. Lyons, R. Pemantle, and Y. Peres. Conceptual proofs of l log l criteria for mean behavior of branching processes. *The Annals of Probability*, 23(3):1125–1138, 1995.

[32] T. F. Móri. On random trees. *Studia Sci. Math. Hungar.*, 39(1-2):143–155, 2002.

[33] T. F. Móri. The maximum degree of the Barabási-Albert random tree. *Comb. Probab. Computing*, 14:339–348, 2005.

[34] T. F. Móri. A surprising property of the Barabási-Albert random tree. *Studia Sci. Math. Hungar.*, 43:265–273, 2006.

[35] Olle Nerman. On the convergence of supercritical general (C-M-J) branching processes. *Z. Wahrsch. Verw. Gebiete*, 57(3):365–395, 1981.

[36] Roberto Oliveira and Joel Spencer. Connectivity transitions in networks with super-linear preferential attachment. *Internet Math.*, 2(2):121–163, 2005.

[37] Peter Olofsson. The $x \log x$ condition for general branching processes. *J. Appl. Probab.*, 35(3):537–544, 1998.

[38] Boris Pittel. Note on the heights of random recursive trees and random m-ary search trees. *Random Struct. Alg.*, 5(2):337–347, 1994.

[39] Anna Rudas. Random tree growth with general weight function. 2004. Arxiv preprint, `http://arxiv.org/abs/math/0410532`.

[40] Anna Rudas and Bálint Tóth. Random tree growth with branching processes - a survey. In B Bollobás, R Kozma, and D Miklós, editors, *Handbook of Large-Scale Random Networks*, volume 18 of *Bolyai Society Mathematical Studies*, chapter 4. Springer, 2009.

[41] Anna Rudas, Bálint Tóth, and Benedek Valkó. Random trees and general branching processes. *Random Struct. Algorithms*, 31(2):186–202, 2007.

[42] Anna Rudas and Imre Péter Tóth. Entropy and Hausdorff dimension in random growing trees. *Stochastics and Dynamics*, Accepted for publication: 2011. 12. 15. DOI Number: 10.1142/S0219493712500104.

[43] Henry Scheffé. A useful convergence theorem for probability distributions. *Ann. Math. Statist.*, 18(3):434–438, 1947.

[44] Robert T. Smythe and Hosam M. Mahmoud. A survey of recursive trees. *Teor. Ĭmovīr. Mat. Stat.*, (51):1–29, 1994.

[45] G. Udny Yule. A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F.R.S. *Royal Society of London Philosophical Transactions Series B*, 213:21–87, 1925.

[46] Remco van der Hofstad. *Random Graphs and Complex Networks.* 2011, `http://www.win.tue.nl/~rhofstad/NotesRGCN2011.pdf`. in preparation.