Continuum Cascade Model: Branching Random Walk for Traveling Wave

Yoshiaki Itoh *

The Institute of Statistical Mathematics and the Graduate University for Advanced Studies

January 5, 2017

Abstract

The cascade model generates random food webs. The continuum cascade model is a Poisson approximation of the cascade model. We have a simple recursion to obtain probability distribution of the longest chain length (the height) generated by the continuum cascade model. Assuming the traveling wave solution, the velocity selection principle for the Fisher-KPP equation works to the recursion.

Here we formulate the height of continuum cascade model as the first passage time of the left most particle of a branching Poisson point process. The asymptotic probability distribution of the height is obtained mathematically by a straightforward application of the Aidekon theorem for the left most particle of branching Poisson point process.

1 Introduction

Random graphs are modeling numerous natural phenomena, polymerization, spread of infectious diseases, transportation systems, electrical distribution

*Institute of Statistical Mathematics, 10-3 Midori-cho, Tachikawa, Tokyo, 190-8562 Japan, Email: itoh@ism.ac.jp
systems, the Internet, the world-wide web, social networks, etc. (see [23] and references therein). A random graph is a set of vertices that are connected by random links. In random graph modeling, links are usually treated as undirected. However, directionality plays a prominent role, in modeling of the web growth, modeling of food webs, etc. The cascade model ([11, 10]) is introduced to analyze the ecological data for community food webs and is a standard model in the subject. The model generates a food web at random. In it the species are labeled from 1 to \( n \), and arcs are given at random between pairs of the species. For an arc with endpoints \( i \) and \( j \) (\( i < j \)), the species \( i \) is eaten by a species labeled \( j \). The longest chain length (the height) \( L_n \) is compared with the ecological data [11]. The cascade model provides a natural mechanism for generating directed random graphs and the same model has been suggested in other contexts, e.g. as a model of parallel computation [17] where the presence of the directed link \((i, j)\) with \( i < j \) indicates that task \( i \) must be performed before task \( j \). For a parallel computation in which each task takes a unit of time, the processing time will be equal to \( M_n + 1 \) for the length of the longest path \( M_n \).

The continuum cascade model is a Poisson approximation of the cascade model [23]. We have a simple recursion to obtain probability distribution of the longest chain length (the height) generated by the continuum cascade model. We applied the velocity selection principle [24] for the Fisher-KPP equation to it assuming the traveling wave solution [23]. The Fisher-KPP equation [15, 24], which is for the spacial spread of an advantageous gene, has a traveling wave solution with the velocity determined by the velocity selection principle. The branching Brownian motion gives a solution to the Fisher-KPP equation and gives the velocity of the wave front with the correction term [7, 26, 27, 34].

Here we obtain the recursion for the probability distribution of the height of continuum cascade model [23] from the position of the first passage time of the branching Poisson point process. We have the asymptotic probability distribution of the height mathematically by a straightforward application of the theorem [4, 5] on the left most particle of the branching random walk [2, 20]. The Aidekon theorem is the branching random walk version of the Lalley and Selke theorem for branching Brownian motion [26] and described in terms of a functional of the limit of the derivative martingale associated to the branching random walk.

Our recursion (2), given in [23], seems to be very typical like the Fisher-
KPP equation. Actually it is pointed out that our continuum cascade model
is also studied under the name of Poisson weighted infinite tree (PWIT, [1, 3,
16]). The closure of vertex 1 in the cascade model converges in distribution to
the PWIT as \( n \) tends to infinity [16]. The expected position of the wave front
and the finite width of the wave front for our continuum cascade is obtained
for the the PWIT mathematically [1], independently from the mathematical
study [4] and independently from our intuitive physical study [23].

Extending the argument on 1-dimensional random sequential packing
([30]), we have the random sequential bisection of intervals [33] (continuum
binary search tree), which has an analogous asymptotic behavior to the bi-
nary search tree of \( n \) keys for the sorting algorithms [13, 31]. The nonlinear
recursion for the probability distribution of the minimum of gaps generated
by 1-dimensional random sequential packing [21, 22] is developed to the non-
linear recursion for the probability distribution [31, 33] of the height of binary
search tree. Assuming the traveling wave solution for the Hattori and Ochiai
conjecture [18, 19, 25], the correction term in the asymptotic expected height
of continuum binary search tree is obtained in [25]. The correction term
is also obtained mathematically [14, 29, 32] without using the assumption.
Like the case of binary search tree, making continuum model gives a natural
mathematical arguments for the original cascade model.

2 A recursion for the continuum cascade model

In the cascade model, the random directed graph has vertex set \( \{1, ..., n\} \n\) in which the directed edges \((i, j)\) occur independently with probability \( c \) for
\( i < j \) and probability zero for \( i \geq j \). Let \( L_n \) denote the length of the
longest path starting from vertex 1. We apply the Poisson approximation to
the binomial distribution of the number of directed edges at each vertex and
consider the continuum cascade model. We study the probability distribution
of \( L_n \) as \( n \) tends to infinity. [10] [28].

At step 1 we generate \( N_x \) points by the Poisson distribution \( Pr(N = k) = \frac{x^k}{k!}e^{-x} \) on \([0, x]\). Each point is mutually independently distributed uni-
formly at random on \([0, x]\). At step \( j (> 1) \), for each generated point at \( x - y \),
generated at the step \( j - 1 \), generate \( N_y \) points by the Poisson distribution
\( Pr(N = k) = \frac{x^k}{k!}e^{-y} \) uniformly at random on the interval \([x - y, x]\), indepen-
dently from the points on other intervals at step \( j \) and independently from
the points of previously generated intervals. We call the terminal interval any interval which did not generate any point. The terminal interval, which appeared at a step, remains a terminal interval after the step. We continue the steps as long as we have at least one interval which is not a terminal interval. For each terminal interval generated by the above procedure, we count the number of steps to get the interval. Let us call the maximum of the numbers, $H(x)$, the height of the tree generated by the continuum cascade model on $[0, x]$.

When $k$ points, $x - y_1, x - y_2, ... , x - y_k$ are generated at step 1, the probability, that the height is not larger than $n - 1$, is $P_{n-1}(y_1)P_{n-1}(y_2)\cdots P_{n-1}(y_k)$. Since each $y_i$ is distributed uniformly at random on $[0, x]$ and $k$ is distributed by the Poisson distribution, we have the following recursion [23] for the probability $P_n(x) \equiv P(H(x) \leq n)$.

For $n = 0$,

$$P_n(x) = e^{-x},$$  \hspace{1cm} (1)

while for $n \geq 1$,

$$P_n(x) = e^{-x} + \sum_{k=1}^{\infty} \frac{x^k}{k!} e^{-x} \frac{1}{x^k} \int_0^x \cdots \int_0^x P_{n-1}(y_1) \cdots P_{n-1}(y_k) \ dy_1 \cdots dy_k$$

$$= e^{-x} \sum_{k=0}^{\infty} \frac{1}{k!} \left( \int_0^x P_{n-1}(y) \ dy \right)^k$$

$$= \exp \left[ -x + \int_0^x P_{n-1}(y) \ dy \right].$$  \hspace{1cm} (2)

# 3 Traveling wave

We apply the Aidekon theorem [4, 5] for branching random walk to show mathematically the following observations [23] for equation (2) in later sections. It is convenient to think about $x$ and $n$ as space and time coordinates, so that the front of the traveling wave was advancing.

1. **Numerical traveling wave solution.** Numerically, the probability distribution $P_n(x)$ has a traveling wave shape with the width of the front remaining finite as shown in Fig. 1, which means the finite width of the height distribution.
Figure 1: Traveling waves, the distribution $P_n(x)$ versus $x$ obtained by iterating the recursion. $P_n(x)$ is shown for $n = 20, 40, 60, 80, 100$ (left to right). Iterations were performed for a discrete approximation (29) given in section 6.

2. **Velocity selection.** Assuming an approximation by a traveling wave form for larger $n \gg 1$,

$$P_n(x) \to \Pi(x - x_f),$$

with the front position $x_f$ growing linearly with ‘velocity’ equal to $e^{-1}$ [23]:

$$x_f \simeq vn, \quad v = \frac{1}{e}. \quad (4)$$

The velocity selection principle [24] gives $v = e^{-1}$ [23] by using an analogous argument to the case of binary search tree [25]. We see that the wave front $x_f$ should advance asymptotically by a constant velocity $v = e^{-1}$, from the probabilistic argument for the cascade model [10, 28].

3. **Logarithmic correction.** An analogy [7, 34] from the Fisher-KPP equation gives a logarithmic correction to the front position as

$$x_f = \frac{n}{e} + \frac{3}{2e} \ln n + O(1). \quad (5)$$

It is convenient to think about $x$ and $n$ as space and time coordinates, so that the front of the traveling wave was advancing.
4 Branching random walk for the solution of the recursion

We consider the asymptotic behavior of branching random walk \([4, 5, 2, 8, 12]\) \([20]\). We follow the notation and argument by Aidekon \([4]\). The process starts with one particle located at 0. At time 1, the particle dies and gives birth to a point process \(L\). Then, at each time \(n \in \mathbb{N}\), the particles of generation \(n\) die and give birth to independent copies of the point process \(L\), translated to their position. If \(T\) is the genealogical tree of the process, we see that \(T\) is a Galton-Watson tree, and we denote by \(|x|\) the generation of the vertex \(x \in T\) (the ancestor is the only particle at generation 0). For each \(x \in T\), we denote by \(V(x) \in \mathbb{R}\) its position on the real line. With this notation, \((V(x); |x| = 1)\) is distributed as \(L\). The collection of positions \((V(x); x \in T)\) defines our branching random walk.

We assume that we are in the boundary case \([20]\)

\[
E[\sum_{|x|=1} 1] > 1, \tag{6}
\]

\[
E[\sum_{|x|=1} e^{-V(x)}] = 1, \tag{7}
\]

\[
E[\sum_{|x|=1} V(x)e^{-V(x)}] = 0. \tag{8}
\]

Every branching random walk satisfying mild assumptions can be reduced to this case by some renormalization. Notice that we may have

\[
\sum_{|x|=1} 1 = \infty \tag{9}
\]

with positive probability \([5]\). We are interested in the minimum at time \(n\)

\[
M_n := \min\{V(x); |x| = n\}, \tag{10}
\]

where \(\min \emptyset = \infty\). Writing for \(y \in \mathbb{R} \cup \{\pm \infty\}\), \(y_+ := \max(y, 0)\), we introduce the random variable

\[
X := \sum_{|x|=1} e^{-V(x)}, \tag{11}
\]

\[
\tilde{X} := \sum_{|x|=1} V(x)_+ e^{-V(x)}. \tag{12}
\]
We assume that the distribution of $L$ is non-lattice, we have

$$E\left[\sum_{|x|=1} V(x)^2 e^{-V(x)}\right] < \infty$$

(13)

$$E[X(\ln_+ X)^2] < \infty, \quad E[\tilde{X}(\ln_+ \tilde{X})] < \infty$$

(14)

To state the result, we introduce the derivative martingale, defined for any $n > 0$ by

$$D_n := \sum_{|x|=n} V(x)e^{-V(x)}.$$  

(15)

From [5, 6] (Proposition A.3 in the Appendix [5]), we know that the martingale converges almost surely to some limit $D_\infty$, which is strictly positive on the set of non-extinction of $T$. Notice that under conditions (6), (7), (8), the tree $T$ has a positive probability to survive.

There exists a constant $C^* > 0$ such that for any real $x$,

$$\lim_{n \to \infty} P(M_n > \frac{3\ln n}{2} + x) = E[e^{-C^*e^x D_\infty}],$$

(16)

(Theorem 1 in [4, 5], see [9] for an elementary approach).

## 5 Probability on the longest chain length

The process starts with one particle located at 0. At time 1, the particle dies and gives birth to the point process $L$, with intensity 1 on $[0, \infty)$. Then, at each time $n \in N$, the particles of generation $n$ die and give birth to independent copies of the point process $L$, translated to their position. At each time, we kill all particles to the right of $x$. Denote position of left-most particle in this (extended) tree at $n$-th generation by $H_n$. Since

$$P(H(x) \leq n - 1) = P(H_n \geq x),$$

(17)

we see

$$P_{n-1}(x) = P(H_n \geq x).$$

(18)
To normalize for equation (8) we replace the original Poisson Point Process of intensity 1 on \([0, \infty)\) by the Poisson Point Process of intensity \(\frac{1}{e}\) on \([-1, \infty)\), as \(\mathcal{L}\) in section 4. Then the conditions (6), (7), (8) hold. We see the inequality (6), since expected number of children here is infinite. For the identity (7) we have

\[
\int_{-1}^{\infty} e^{-y} \frac{dy}{e} = 1
\]

and for the identity (8) we have

\[
\int_{-1}^{\infty} ye^{-y} \frac{dy}{e} = 0.
\]

The distribution of this Poisson point process \(\mathcal{L}\) is non-lattice and the moment conditions (13) and (14) hold by exponential decay (7) for \(V(x)\). We have

\[
E\left[ \sum_{|x|=1} V(x)^2 e^{-V(x)} \right] = \int_{-1}^{\infty} x^2 e^{-x} dx < \infty.
\]

The total number of children is assumed to be finite almost surely in [4]. However the argument [4] is applied to the above extension to \([-1, \infty)\), as shown in [5].

The position \(H_n\) of the left-most particle at generation \(n\) is given by using \(M_n\) for

\[
M_n = eH_n - n.\tag{22}
\]

Considering equation (22),

\[
M_n > z + \frac{3}{2} \ln n,\tag{23}
\]

means

\[
H_n > \frac{z + n + \frac{3}{2} \ln n}{e}.\tag{24}
\]
Hence from equation (18),
\[ P(M_n > z + \frac{3\ln n}{2}) = P(H_n > \frac{z + n + \frac{3}{2}\ln n}{e}) \]
\[ = P_{n-1}(\frac{z + n + \frac{3}{2}\ln n}{e}) \quad (25) \]

Put \( z/e = x \), then from equation (16) (Aidekon [4] [5]), for the solution \( P_{n-1} \) to equation (2) we have
\[ \lim_{n \to \infty} P_{n-1}(x + \frac{n}{e} + \frac{3}{2e}\ln n) = E[\exp(-C^*e^{ex}D_{\infty})], \]
\[ (27) \]
which gives the asymptotic probability on the longest chain length (on the position of wave front).

For \( x = 0 \) of equation (27), we have
\[ \lim_{n \to \infty} P_{n-1}(\frac{n}{e} + \frac{3}{2e}\ln n) = E[\exp(-C^*D_{\infty})], \]
\[ (28) \]
which should be less than 1 and larger than 0, since \( D_{\infty} \) is mathematically shown to be strictly positive [4, 5, 6].
6 Numerical observation

Putting $\bar{x}\Delta$ for $x$, and giving the discrete initial value for equation (1), we consider a recursion as a discretization of equation (2) for $1 \leq n$,

$$f_n(\bar{x}) = \exp[-\bar{x}\Delta + \sum_{\bar{y}=1}^\Delta f_{n-1}(\bar{y})\Delta].$$

(29)

The numerical value $f_n(\bar{x})$ for $P_n(x)$ in Fig. 1 and Fig. 2 are obtained from equation (29) for $\Delta = 0.01$ by using the software Mathematica. The numerical values

$$f_{n-1}\left(\frac{1}{\Delta}\left(\frac{3}{2e} \ln n + \frac{n}{e}\right)\right)$$

(30)

for $P_{n-1}(\frac{3}{2e} \ln n + \frac{n}{e})$, are shown by the lower curve in Fig. 2. Putting $e/\alpha$ instead of exponential $e$, the numerical values

$$f_{n-1}\left(\frac{\alpha}{\Delta}\left(\frac{3}{2e} \ln n + \frac{n}{e}\right)\right)$$

(31)

for $\Delta = 0.01$ and $\alpha = 0.9855$ are shown by the upper curve in Fig. 2, which seems to approach quickly to a constant. We carried out calculations and see for example $\Delta = 0.001$ and $\alpha = 0.9977$ the value of (31) quickly approaches to a constant. Our numerical calculations seem to suggest $\alpha \to 1$ as $\Delta \to 0$, which supports equation (28) for the wavefront numerically.

Acknowledgements The author thanks Amir Dembo and Ofer Zeitouni for suggesting to him an essential idea given in section 5. The author thanks Elie Aidekon for the reply [5] to the question of Amir Dembo on [4], which is essential for the present paper. This work is supported in part by US National Science Foundation Grant DMS1225529 to Rockefeller University and JSPS Grant-in-aid for Scientific Research 23540177.

References

[1] L. Addario-Berry and K. Ford, Poisson-Dirichlet branching random walks, The Annals of Applied Probability 23, 283-307 (2013).
[2] L. Addario-Berry and B. Reed, Minima in branching random walks. The Annals of Probability, 37, 1044-1079 (2009).

[3] D. Aldous, and J. M. Steele, The objective method: probabilistic combinatorial optimization and local weak convergence, Probability on discrete structures, Springer Berlin Heidelberg, 1-72 (2004).

[4] E. Aidekon, Convergence in law of the minimum of a branching random walk, The Annals of Probability, 41, 1362-1426 (2013).

[5] E. Aidekon, Convergence in law of the minimum of a branching random walk, http://arxiv.org/pdf/1101.1810v6.pdf (2013).

[6] J. D. Biggins and A. E. Kyprianou, Measure change in multitype branching, Advances in Applied Probability, 36, 544-581 (2004).

[7] M. Bramson, Maximal displacement of branching Brownian motion, Communications on Pure and Applied Mathematics 31. 531-581 (1978).

[8] M. Bramson and O. Zeitouni, Tightness for a family of recursion equations, The Annals of Probability, 37, 615-653 (2009).

[9] M. Bramson, J. Ding and O. Zeitouni, Convergence in law of the maximum of nonlattice branching random walk, http://arxiv.org/abs/1404.3423 (2014).

[10] J. E. Cohen and C. M. Newman, A stochastic theory of community food webs I. Models and aggregated data, Proceedings of the Royal Society of London B 224, 421–448 (1985).

[11] J. E. Cohen, F. Briand and C. M. Newman, Community food webs: data and theory. Vol. 20. Springer Science & Business Media, (2012).

[12] A. Dembo and O. Zeitouni, Large deviations techniques and applications. New York: Springer (1998).

[13] L. Devroye, A note on the height of binary search trees, Journal of the ACM (JACM), 33, 489–498 (1986).

[14] M. Drmota, An analytic approach to the height of binary search trees II, Journal of the ACM (JACM), 50, 333-374 (2003).
[15] R. A. Fisher, The wave of advance of advantageous genes, Annals of Eugenics, 7, 355-369 (1937).

[16] K. Gabrysch, Convergence of directed random graphs to the Poisson-weighted infinite tree, Journal of Applied Probability 53, 463-474 (2016).

[17] E. Gelenbe, R. Nelson, T. Philips and A. Tantawi, An approximation of the processing time for a random graph model of parallel computation, ACM 86 Proceedings of 1986 ACM Fall joint computer conference IEEE Computer Society Press, Los Alamos, 691–697 (1986).

[18] T. Hattori and H. Ochiai, Scaling limit of successive approximations for w=w² and its consequences on the theories of random sequential bisections and height of binary search trees, preprint (1998). (http://web.econ.keio.ac.jp/staff/hattori/200411.pdf)

[19] T. Hattori and H. Ochiai, Scaling Limit of Successive Approximations for w’=-w². Funkcialaj Ekvacioj, 49, 291-319 (2006).

[20] Y. Hu and Z. Shi, Minimal position and critical martingale convergence in branching random walks, and directed polymers on disordered trees. The Annals of Probability, 37, 742-789 (2009).

[21] Y. Itoh, Random Packing Model for Nomination of Candidates and its Application to Elections of the House of Representatives in Japan, Proceedings of the International Conference on Cybernetics and Society, IEEE (1978) 432–435.

[22] Y. Itoh, On the minimum of gaps generated by one-dimensional random packing, Journal of Applied Probability, 17, 134-144 (1980).

[23] Y. Itoh and P. L. Krapivsky, Continuum cascade model of directed random graphs: traveling wave analysis, Journal of Physics A: Mathematical and Theoretical, 45, 455002 (2012).

[24] A. N. Kolmogorov, I. G. Petrovskii and N. S. Piskunov, A study of the equation of diffusion with increase in the quantity of matter, and its application to a biological problem. Bjul. Moskovskogo Gos. Univ, 1(7), 1-26 (1937).
[25] P. L. Krapivsky and S. N. Majumdar, Traveling Waves, Front selection, and exact nontrivial exponents in a random fragmentation problem, Physical Review Letters, 85, 5492–5495 (2000).

[26] S. P. Lalley and T. Sellke, A conditional limit theorem for the frontier of branching Brownian motion, Annals of Probability 15, 1052-1061 (1987).

[27] H. P. McKean, Application of Brownian motion to the equation of Kolmogorov-Petrovskii-Piskunov, Communications on Pure and Applied Mathematics, 28, 323-331 (1975).

[28] C. M. Newman, Chain lengths in certain random directed graphs, Random Structure and Algorithms, 3, 243–253 (1992).

[29] B. Reed, The height of a random binary search tree. Journal of the ACM (JACM), 50, 306-332 (2003).

[30] A. Renyi, On a one-dimensional problem concerning random space filling, Publ. Math. Inst. Hung. Acad. Sci. 3.109-127 (1958).

[31] J. M. Robson, The height of binary search trees, Australian Computer Journal, 11, 151–153 (1979).

[32] L. Shepp, D. Zeilberger and C. H. Zhang, Pick up sticks, arXiv preprint arXiv:1210.5642 (2012).

[33] M. Sibuya and Y. Itoh, Random sequential bisection and its associated binary tree, Annals of the Institute of Statistical Mathematics, 39, 69-84 (1987).

[34] K. Uchiyama, The behavior of solutions of some non-linear diffusion equations for large time, Journal of Mathematics of Kyoto University, 18, 453-508 (1978).