Simulation results on a triangle-based network evolution model[*]

István Fazekas, Attila Barta, Csaba Noszály

University of Debrecen
`fazekas.istvan@inf.unideb.hu`
`barta.attila@inf.unideb.hu`
`noszaly.csaba@inf.unideb.hu`

Abstract

We study a continuous time network evolution model. We consider the collaboration of three individuals. In our model, it is described by three connected vertices, that is by a triangle. During the evolution new collaborations, that is new triangles are created. The reproduction of the triangles is governed by a continuous time branching process. The long time behaviour of the number of triangles, edges and vertices is described. In this paper, we highlight the asymptotic behaviour of the network by simulation results.

Keywords: random graph, network, branching process, Malthusian parameter

MSC: 05C80, 90B15, 60J85

1. Introduction

In the past two decades network science became a popular and important topic, see [2]. It describes large real-life networks as the Internet, the WWW, social, biological and energy networks. Large networks have several common properties, therefore it is worth to study theoretical models of networks. Usually, networks are described by graphs. The nodes of the network are the vertices of the graph

and the connections are the edges. The meaning of connection can be cooperation or any interaction. A most cited paper in network science is [3]. It studies the famous preferential attachment model which leads to scale free networks. A deep mathematical study of discrete time network evolution models can be found in [6].

However, in our paper, we turn to a continuous time network evolution model. An interesting continuous time model is presented in [7]. In that paper the theory of general branching processes, so called Crump-Mode-Jagers processes (see [8]) is applied to obtain asymptotic theorems. In paper [9], the idea of preferential attachment is combined with the evolution mechanism of a multi-type continuous time branching process.

In this paper we apply certain ideas of papers [1] and [7]. Paper [1] describes the interaction (or co-operation) of three persons. It creates a discrete time network evolution model which relies on the preferential attachment rule and three-interactions. In [7], however, a continuous time network evolution model based on a branching process evolution rule is presented. In that model only the usual interaction of two vertices is included and triangles have no role in the evolution rule. Neither [1] nor [7] offer numerical results. In our paper we combine the above ideas of three interactions with the continuous time branching process evolution mechanism. We focus on numerical studies of our model.

In this paper, in Section 2, we offer a detailed description of the evolution rules of our network. Then, in Section 3, we give a brief summary of our theoretical results. Their detailed mathematical proofs are given in a separate paper (see [5]). Here, in Section 4, we present our numerical results. We show that our formulae are numerically tractable, so we can calculate the values of the important parameters and other features of our process. Then we show our simulation program and a certain part of our simulation results. These results support our mathematical theorems.

2. The network evolution model

We shall study the following evolving random graph model. At the initial time $t = 0$ we start with a single triangle. During the evolution new triangles are born. Every triangle has its own evolution process. We assume that during the evolution of the network the life processes of the triangles are identically distributed and independent of each other.

We denote the reproduction process of the generic triangle by $\xi(t)$ and its birth times by $\tau_1, \tau_2, \ldots$. We assume that $\tau_1, \tau_2, \ldots$ are the jumping time points of a Poisson process $\Pi(t), t \geq 0$, where the rate of Π is equal to 1. Then the point process $\xi(t)$ gives the total number of offspring up to time t. However, at a birth time not only triangles can be created but new edges and vertices can be added to the graph. Here we describe the details of an evolution step. At every birth time τ_i a new vertex is added to the graph which can be connected to its ancestor triangle with j edges, where $j = 0, 1, 2, 3$. The vertices of the ancestor triangle to be connected to the new vertex are chosen uniformly at random. Let p_j denote the probability that

the new vertex will be connected to j vertices of the ancestor triangle. It follows from the definition of the evolution process that at each birth step the possible number of the new triangles can be $0, 1$ or 3. On Figure 1 we represent these possibilities. The initial triangle is drawn by solid lines while the new ingredients by dashed lines. Denote the litter sizes belonging to the birth times $\tau_1, \tau_2, \ldots$ by
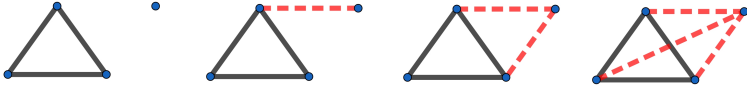

Figure 1: Possible birth events (0, 1, 2 or 3 new edges)

$\varepsilon_1, \varepsilon_2, \ldots$. Then $\varepsilon_1, \varepsilon_2, \ldots$ are independent identically distributed discrete random variables with distribution $\mathbb{P}(\varepsilon_i = j) = q_j, j \geq 0$. In our model the distribution of the litter size ε_i is given by

$$\mathbb{P}(\varepsilon_i = 0) = q_0 = p_0 + p_1, \ \mathbb{P}(\varepsilon_i = 1) = q_1 = p_2, \ \mathbb{P}(\varepsilon_i = 3) = q_3 = p_3,$$

$$\mathbb{P}(\varepsilon_i = j) = q_j = 0, \ \text{if} \ j \notin \{0, 1, 3\}.$$

We assume that the litter sizes are independent of the birth times $\tau_1, \tau_2, \ldots$, too. Denote by λ the life length of the generic triangle. λ is a finite nonnegative random variable. After its death the triangle does not produce offspring, therefore $\xi(t) = \xi(\lambda)$ when $t > \lambda$. Then the reproduction process of a triangle is

$$\xi(t) = \sum_{\tau_i \leq t \wedge \lambda} \varepsilon_i = S_{\Pi(t \wedge \lambda)},$$

where $S_n = \varepsilon_1 + \cdots + \varepsilon_n$ gives the total number of offspring before the $(n+1)$th birth event and $x \wedge y$ denotes the minimum of $\{x, y\}$.

Let $L(t)$ be the distribution function of λ. We assume that the survival function of the triangle's life length is

$$1 - L(t) = \mathbb{P}(\lambda > t) = \exp\left(-\int_0^t l(u)\, \mathrm{d}u\right),$$

where $l(t)$ is the hazard rate of the life span λ. Moreover, we assume that the hazard rate depends on the number of offspring as

$$l(t) = b + c\xi(t)$$

with positive constants b and c.

The whole evolution process is the following. The life and the reproduction process of the initial triangle is the same as that of the above described generic triangle. When a child triangle is born, then it starts its own life and reproduction process which is also defined by the same way as its parent triangle. The same applies to

the grandchildren, etc. Therefore the evolution of the network is described by a continuous time branching process. We underline that the life and reproduction process of any triangle have the same distribution as those of the generic triangle, but the reproduction processes of different triangles are independent.

3. Theoretical results

Here we summarize the theoretical results of our paper [5].

Let $\mu(t) = \mathbb{E}\xi(t)$ be the expectation of the number of offspring of a triangle up to time t. The total number of offspring of a triangle is $\xi(\infty)$. The expected offspring number of a triangle can be calculated as

$$\mu(\infty) = \mathbb{E}\xi(\infty) = (q_1 + 3q_3)\mathbb{E}(\lambda) =$$

$$(q_1 + 3q_3)\frac{1}{c}\int\limits_0^1 (1-u)^{\frac{b+1-q_0}{c}-1} e^{\frac{u}{3c}(q_3 u^2 - 3q_3 u + 3(q_1+q_3))}\, du.$$

The probability of extinction is 1 if $\mu(\infty) \leq 1$.

Theorem 3.1. *If $\mu(\infty) > 1$, then the probability of the extinction of the triangles is the smallest non-negative solution of equation*

$$\frac{q_1 + q_3(y^2 + y + 1)}{c}\int\limits_0^1 (1-u)^{\frac{1+b-q_0}{c}-1} e^{(\frac{q_1 y + q_3 y^3}{c}u - \frac{q_3 y^3}{c}u^2 + \frac{q_3 y^3}{3c}u^3)}\, du = 1. \quad (3.1)$$

Assume that $\mu(\infty) > 1$, that is our branching process is supercritical. Then the Malthusian parameter α is the only positive solution of equation $\int_0^\infty e^{-\alpha t}\mu(dt) = 1$. We can see that

$$q_1 + 3q_3 - b - 1 < \alpha < q_1 + 3q_3 - b.$$

In our model the Malthusian parameter α satisfies the equation

$$1 = \frac{(q_1 + 3q_3)}{c}\int\limits_0^1 (1-u)^{\frac{\alpha+(b+1)}{c}-\frac{q_0}{c}-1} e^{\frac{3q_1 u + q_3 u(u^2 - 3u + 3)}{3c}}\, du. \quad (3.2)$$

Now we give the asymptotic behaviour of the number of triangles. Let us denote by $Z(t)$ the number of triangles alive at time t. Let α be the Malthusian parameter.

Theorem 3.2. *We have*

$$\lim_{t\to\infty} e^{-\alpha t}Z(t) = Y_\infty m_\infty$$

almost surely and in L^1, where the random variable Y_∞ is nonnegative, it is positive on the event of non-extinction, it has expectation 1 and

$$m_\infty = \frac{1}{(q_1 + 3q_3)^2 \int_0^\infty t e^{-\alpha t}(1 - L(t))dt}.$$

Now we turn to the asymptotic behaviour of vertices and edges. Let us denote by $V(t)$ the total number of vertices (dead or alive) up to time t. Let $W(t)$ be the number of edges (dead or alive) up to time t. Let γ denote the number of new edges at a birth. Then its distribution is $\mathbb{P}(\gamma = j) = p_j$, $j = 0, 1, 2, 3$.

Theorem 3.3. *We have*

$$\frac{V(t)}{Z(t)} \to \frac{1}{\alpha} \qquad and \qquad \frac{W(t)}{Z(t)} \to \frac{\mathbb{E}\gamma}{\alpha}$$

as $t \to \infty$ almost surely on the event of non-extinction.

4. Numerical and simulation results

To get a closer look on the theoretical results, we made some simulations about them. We generated our code in Julia language [4]. We chose Julia, because of the great implementation of priority queues. The simulation time of our code was significantly faster in Julia than in other programming languages. We handled the main objects (the triangles) of our model as arrays with 3 elements. The elements were the indices of the edges that formed an individual for the process. We put all triangles in a priority queue with the priority of its birth time, because we can pop out the element with the lowest priority. After we have got the triangle with the lowest birth time, we can handle its birth process with the predefined parameters b, c, q_1, q_3. In the birth process we generated an exponential time step for the next birth step of our triangle. After that we checked if the triangle is still alive by calculating the survival function. If the triangle is dead, we move to the next one. If it is alive, then we generate 1 or 3 new triangles and put them into the priority queue with the calculated birth time priorities. After this step we moved to the next birth event. The pseudocode of the birth process is seen at Algorithm 1.

We made several simulation experiments. Here we show only some typical results. For the above demonstration first we used the parameter set $b = 0.2$, $c = 0.2$, $p_0 = 0.05$, $p_1 = 0.05$, $p_2 = q_1 = 0.6$, $p_3 = q_3 = 0.3$. On Figure 2a the solid curve shows the number of triangles. According to Theorem 3.2 it has asymptotic rate $e^{\alpha t}$. Therefore we put logarithmic scale on the vertical axis, so the function $Z(t)$ is a straight line for large values of t. On the figure one can see that the shape of the curve is close to a straight line, so it supports our Theorem 3.2.

Then we checked the value of the Malthusian parameter α. We can find it in two ways. On the one hand, the slope of the line on Figure 2a is α for large values of the time. This slope can be approximated by the differences of the function. So on Figure 2b we present these differences (solid line). On the other hand, α can be calculated numerically from equation (3.2). This α value is shown of Figure 2b by the horizontal dashed line. The fit of the differences to this α can be seen for large values of t. To get a closer look on the Malthusian parameter α, we fixed 5 parameter sets. Then we calculated α from equation (3.2) for each case. It is shown in the fifth column of Table 1. Then for each of the parameter sets we

(a) The number of triangles in
the evolving graph

(b) The differences of
the number of triangles
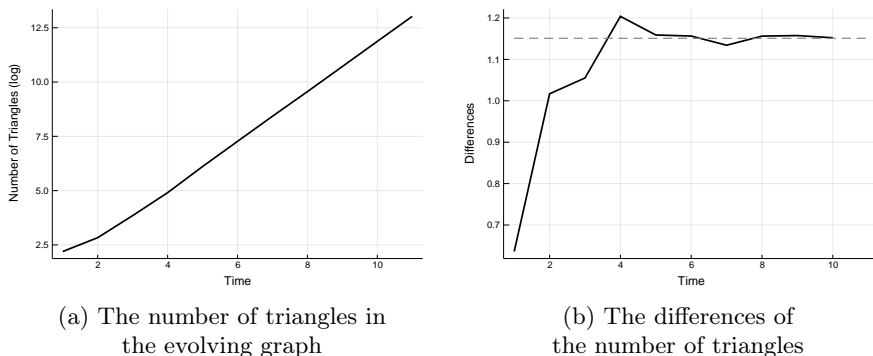
Figure 2: Simulation results for $b = 0.2$, $c = 0.2$, $q_1 = 0.6$, $q_3 = 0.3$

simulated our process $Z(t)$ five times. Then we calculated the differences of $\log Z(t)$ which should be good approximations of α according to Theorem 3.2. In Table 1, $\widehat{\alpha_1}$, $\widehat{\alpha_2}$, $\widehat{\alpha_3}$, $\widehat{\alpha_4}$, $\widehat{\alpha_5}$ show the values of these approximations for large t. One can see that each $\widehat{\alpha_i}$ is close to the corresponding α. We calculated numerically the

b	c	q_1	q_3	α	$\widehat{\alpha_1}$	$\widehat{\alpha_2}$	$\widehat{\alpha_3}$	$\widehat{\alpha_4}$	$\widehat{\alpha_5}$
0.2	0.4	0.7	0.1	0.5628	0.5651	0.5730	0.5701	0.5611	0.5594
0.2	0.4	0.8	0.1	0.6531	0.6537	0.6497	0.6570	0.6510	0.6589
0.4	0.4	0.8	0.1	0.4531	0.4503	0.4519	0.4584	0.4541	0.4524
0.4	0.4	0.7	0.2	0.6545	0.6533	0.6517	0.6548	0.6534	0.6574
0.4	0.4	0.6	0.3	0.8535	0.8519	0.8489	0.8559	0.8547	0.8566

Table 1: α from equation (3.2) and $\widehat{\alpha_i}$ from simulations

probability of extinction from equation (3.1). It is shown in the column 'Numerical' of Table 2. In the column 'Simulation' the relative frequency of the extinction is shown using our computer experiment. For each parameter sets, we simulated 10^4 processes and counted the number of extinctions occured. The value of the relative frequency is close to the corresponding value of the probability in each case. So Table 2 supports the result of Theorem 3.1. To investigate how our difference process approximates α for large values of time t, we simulated around 500 independent processes with the same parameters $b = 0.2$, $c = 0.2$, $p_0 = 0.05$, $p_1 = 0.05$, $p_2 = q_1 = 0.6$, $p_3 = q_3 = 0.3$ and same running time. Then we checked the differences of the last two values in the numbers of triangles that we simulated and made a histogram, seen in Figure 3. From equation (3.2) we obtained that the value of α is 0.3365. We see that the values of the differences are in $[0.332, 0.340]$, so they are very close to 0.3365.

b	c	q_1	q_3	Simulation	Numerical
0.0	0.2	0.4	0.4	0.0	0.0
0.1	0.2	0.4	0.4	0.1304	0.1282
0.1	0.2	0.5	0.4	0.1165	0.1158
0.1	0.2	0.5	0.5	0.1097	0.1025
0.2	0.2	0.5	0.4	0.2227	0.2180
0.2	0.2	0.6	0.4	0.2038	0.2002
0.3	0.3	0.5	0.4	0.3231	0.3185
0.4	0.4	0.5	0.4	0.3966	0.4020

Table 2: The relative frequency and the probability of the extinction of the triangles
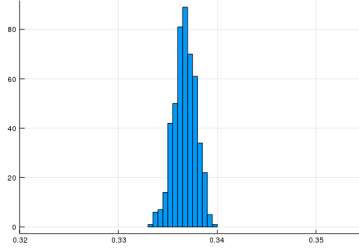

Figure 3: Histogram of differences

To get some information about the random variable $Y_\infty m_\infty$ represented in Theorem 3.2, we calculated the value of $Z(t)e^{-\alpha t}$ for 10^3 independent repetitions of the process for the same time t and same parameters $q_1 = 0.3$, $q_3 = 0.6$, $b = 0.2$, $c = 0.2$. On Figure 4 we represent the histogram and the empirical cumulative distribution function (ECDF) calculated from the simulation. We obtained that the empirical cumulative distribution function of $Y_\infty m_\infty$ fits to a gamma distribution, as the Kolmogorov–Smirnov test gave us p value 0.6713.
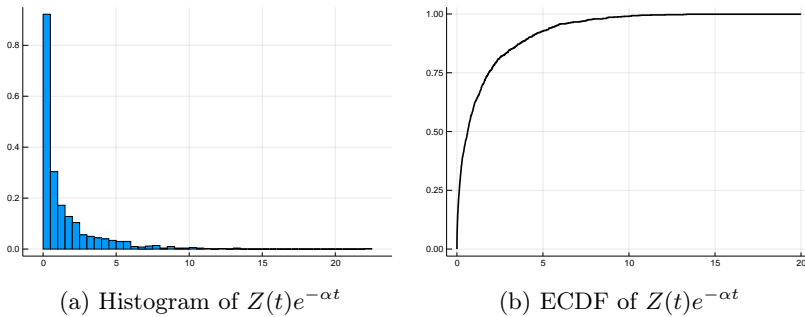

(a) Histogram of $Z(t)e^{-\alpha t}$

(b) ECDF of $Z(t)e^{-\alpha t}$

Figure 4: Simulation results for $Z(t)e^{-\alpha t}$

5. Summary

In this paper and in [5] we offer a network evolution model which describes collaborations of 3 persons. Our model grasps certain features of real collaborations as emerging and disappearing collaborations, moreover collaborations of 2 persons are also allowed. Our numerical results confirm the theoretical ones. The present results prepare our future research on more complicated collaborations.

Algorithm 1 Birth process of a triangle

1: **procedure** BIRTH PROCESS
2:      $Y \leftarrow$ non-empty *Priority Queue*
3:      $b, c, q_1, q_3 \leftarrow$ parameters of the survival function
4:      $x \leftarrow$ dequeue Y
5:      **if** x is a new triangle **then**
6:          $t_0 \leftarrow$ the birth time of x in the whole process
7:          $t \leftarrow 0$, lifetime of x
8:          $l \leftarrow 1$, life variable
9:          **while** $l = 1$ **do**
10:              $t \leftarrow t + Exp(1)$
11:              $p \leftarrow$ the calculated survival function
12:              **if** $p > Uni(0, 1)$ **then**
13:                  $p_0 \leftarrow Uni(0, 1)$
14:                  **if** $p_0 < q_1$ **then**
15:                      take a new triangle with $t_0 + t$ birth time to Y
16:                      offspring number is 1 at birth time t
17:                  **else if** $p_0 > 1 - q_3$ **then**
18:                      take three new triangles with $t_0 + t$ birth times to Y
19:                      offspring number is 3 at birth time t
20:                  **else**
21:                      $l \leftarrow 0$
22:                      take t as the death time of x to Y

References

[1] Á. BACKHAUSZ, T. F. MÓRI: *A random graph model based on 3-interactions*, Ann. Univ. Sci. Budapest 36 (2012), pp. 41–52.

[2] A.-L. BARABÁSI: *Network science*, Cambridge: Cambridge University Press, 2018.

[3] A.-L. BARABÁSI, R. ALBERT: *Emergence of scaling in random networks*, Science 286.5439 (1999), pp. 509–512,
DOI: https://doi.org/10.1126/fscience.286.5439.509.

[4] J. BEZANSON, A. EDELMAN, S. KARPINSKI, V. B. SHAH: *Julia: A Fresh Approach to Numerical Computing*, 2014, arXiv: 1411.1607 [cs.MS].

[5] I. Fazekas, A. Barta, C. Nószály, B. Porvázsnyik: *A continuous time evolution model describing 3-interactions*, in: Manuscript, Debrecen, Hungary, 2020.

[6] R. van der Hofstad: *Random Graphs and Complex Networks Vol. 1.* Cambridge: Cambridge Series in Statistical and Probabilistic Mathematics, 2017, DOI: https://doi.org/10.1017/9781316779422.

[7] T. F. Móri, R. Sándor: *A random graph model driven by timedependent branching dynamics*, Ann. Univ. Sci. Budapest, Sect. Comp. 46.3 (2017), pp. 191–213.

[8] O. Nerman: *On the convergence of supercritical general (C-M-J) branching processes*, Z. Wahrscheinlichkeitstheorie verw Gebiete 57 (1981), pp. 365–395, DOI: https://doi.org/10.1007/BF00534830.

[9] S. Rosengren: *A Multi-type Preferential Attachment Tree*, Internet Math. 2018 (2018).