Short Communication

Geometric preferential attachment with choice-based edge step

Y.A. Malyshkin
Tver State University, Tver Russia
yury.malyshkin@mail.ru

DOI: 10.26456/pcascnn/2024.16.381

Abstract: We study the asymptotic behavior of the maximum degree in the geometric graph model with a preferential attachment choice-based edge step. Geometric graphs are natural models that describe some nanoscale systems, while preferential attachment provides a good description of complex networks, particularly different neural networks. The model is a recursively built sequence of graphs. We start with the initial graph on a single vertex and we add a new vertex and draw a few edges on each step. Each vertex is assigned a parameter that represents its location. The recursion step consists of two parts. First, we introduce a new vertex and draw edges to close enough vertices. This step represents the geometric part of the model. Then, we draw edges between vertices by preferential attachment with the choice rule. We prove that dependent on model parameters, the maximum degree could exhibit sublinear (similar to the standard preferential attachment) and linear (representing concentration effect) behavior.

Keyword: geometric graphs, complex networks, random graphs, preferential attachment, power of choice.

1. Introduction

Random graphs are widely used to model different nanoscale networks and systems (see, e.g., [1]). For example, percolation on geometrical graphs could be used to describe nanostructured materials (see, e.g., [2]) and geometrical power law graphs well describe neural networks (see, e.g., [3]). The last could be used to model different processes in a variety of materials.

In the present work, we study the addition of the choice-based edge step to the geometry attachment model. The standard preferential attachment graph model was introduced in [4]. The general idea behind preferential attachment models is that vertices with higher degrees are more likely to attract edges from newly introduced vertices. Usually, one considers the case when the probability of drawing an edge to a vertex is proportional to its degree's linear function (see, e.g. [5, 6]).

One of the modifications of such a model is introducing a choice to the model (see, e.g., [7-10]). In this modification, we consider the sample of d independently chosen vertices and choose the one with the largest degree. This modification often results in the effect of condensation, when a single vertex has linear (over the total number of edges) degree (see, e.g., [11-13]). The other modification is the introduction of the edge step to the model (see, e.g., [14]).

Let us introduce our model. Fix $m,d \in \mathbb{N}$, $d>1$, $r>0$, $\beta>-1$, which are parameters of our model. We consider a sequence of graphs $\{G_n\}$, $n \in \mathbb{N}$, with G_n containing vertices $v_1,...,v_n$. We consider i.i.d. random variables $\mathrm{X}_1,...,\mathrm{X}_n$

distributed uniformly in [0,1], such that X_i corresponds to v_i and represents the location of a vertex. We start with the initial graph G_1 that consists of a vertex v_1. To build graph G_{n+1} from G_n we add a vertex v_{n+1} with location X_{n+1} and draw edges in two steps.

1. Vertex step: we draw edges from v_{n+1} to vertices with distance at most r/n from it.

2. Edge step: We choose m vertices u_n^i, $i=1,...,m$, uniformly among all vertices. Then, we draw an edge between a pair $\left(u_n^i, w_n^i\right)$, where a vertex w_n^i is chosen by the following rule. We consider a sample $y_n^{i,1},...,y_n^{i,d}$ of size d of vertices of $G_n \cup v_{n+1} - w_n^i$, chosen with probabilities proportional to their degrees plus β. Then w_n^i is the vertex from the sample with the highest degree in G_n (in case of a tie we chose randomly, it would not affect the degree distribution).

2. Results

Let $D(n)=\sum_{j=1}^{n}\left(\deg_{G_n}(v_j)+\beta\right)$ be the total weight of all vertices. At step n+1 we could increase $D(n)$ in the following ways. The new vertex could be at a distance at most r/n from another vertex, for each vertex it happens with probability $2r/n$, so the average increase would be $2r$. Also, with high probability, there could be at most $\ln^2 n$ vertices within any neighborhood of size r. We also increase $D(n)$ during an edge step by drawing m edges.

Therefore, we get

$$\mathbb{E}\left(D(n+1)-D(n)\,|\,\mathcal{F}_n\right)=2m+2r+O\left(\frac{\ln^2 n}{n}\right).$$

By stochastic approximation (see, e.g., Theorem 3.1.1 in [15]),

$$D(n+1)-D(n)=2m+2r+O\left(\frac{1}{n^{2/3}}\right) \tag{1}$$

almost surely.

Let us formulate our main result. Let $M(n)$ be the highest degree of vertices in G_n. Let define function

$$f(x):=m\left(1-\left(1-\frac{x}{2m+2r+\beta}\right)^d\right).$$

Theorem 1. In the model described above

1. If $\dfrac{dm}{2m+2r+\beta}<1$, then, for any $\epsilon>0$:

$$\mathbb{P}\left(\forall n>N: n^{\frac{dm}{2m+2r+\beta}-\epsilon}<M_i(n)<n^{\frac{dm}{2m+2r+\beta}+\epsilon}\right)\to 1$$

as $N \to \infty$.

2. If $\frac{dm}{2m+2r+\beta} > 1$, then almost surely

$$\liminf_{n\to\infty} \frac{M_i(n)}{n} = x^*$$

where x^* is a unique positive root of the equation $f(x) = x$.

Note that the unique root exists due to the concavity of $f(x)$ and the condition $f'(0) > 1$.

To prove the above theorem, we need the following auxiliary result from [16].

Lemma 1. Let $\mathcal{F}_n$-measurable process $Y(n)$ with values in $\mathbb{N}$ with non-negative bounded increments (i.e. $0 \le Y(n+1) - Y(n) \le C$) satisfies

$$\mathbb{E}(Y(n+1) - Y(n) \mid \mathcal{F}_n) = g\left(\frac{Y(n)}{n}\right) + O\left(\frac{1}{n}\right),$$

where $g(x)$ satisfies $g(0) = 0$, $g'(0) > 0$ and $g''(x) < 0$.

1. If $g'(0) < 1$ than for any $\epsilon > 0$

$$\mathbb{P}\left(\forall n > n_0 : n^{g'(0)-\epsilon} < Y(n) < n^{g'(0)+\epsilon}\right) \to 1, \text{ as } n_0 \to \infty.$$

2. If $g'(0) > 1$, then almost surely

$$\lim_{n\to\infty} \frac{Y(n)}{n} = x^*,$$

where x^* is a unique positive root of the equation $g(x) = x$.

3. Proof of the main result

Due to (1), for any $\epsilon > 0$:

$$\mathbb{P}\left(\mathcal{A}_\epsilon(n_0)\right) \to 1 \text{ as } n_0 \to \infty,$$

where $\mathcal{A}_\epsilon(n_0) = \left\{\forall n \ge n_0, \mid D_i(n) - (2m+2r+\beta)n \mid < \epsilon n\right\}$.

Let us consider the evolution of $M(n)$. On step $n+1$ it could be increased in two ways. First, we could put a new vertex at a distance at most r from a vertex with degree $M(n)$, which happens with conditional probability $L(n)/n$, where $L(n)$ is the number of vertices of degree $M(n)$. Second, we could draw edges (the same procedure independently repeated k times) to the vertex of degree $M(n)$ as the second vertex during an edge step (or for it to be the initial vertex, which happens with probability $1/n$). To do so, we need a vertex with the highest degree to appear in the sample. The probability of getting the vertex to the exact position in the sample is $\frac{(M(n)+\beta)L(n)}{D(n)}$, so the probability for the vertex of the maximal degree to be in the sample is:

$$1-\left(1-\frac{(M(n)+\beta)L(n)}{D(n)}\right)^d.$$

Therefore, we get:

$$\mathbb{E}(M(n+1)-M(n)\,|\,\mathcal{F}_n)=m\left(1-\left(1-\frac{(M(n)+\beta)L(n)}{D(n)}\right)^d\right)+O\left(\frac{L(n)}{n}\right).$$

By the persistent hub arguments similar to [16], we get that exists random $N<\infty$, that $L(n)=1$ for all $n>N$.

Hence, we have

$$\mathbb{E}(M_1(n+1)-M_1(n)\,|\,\mathcal{F}_n)=m\left(1-\left(1-\frac{(M(n)+\beta)}{D(n)}\right)^d\right)+O\left(\frac{1}{n}\right)$$

almost surely.

As a result, for $n\geq N$ we get

$$\mathbf{1}\{\mathcal{A}_\epsilon^1(n_0)\}\mathbb{E}(M(n+1)-M(n)\,|\,\mathcal{F}_n)\geq\mathbf{1}\{\mathcal{A}_\epsilon^1(n_0)\}\left(f_{\epsilon,1}^+\left(\frac{M(n)}{n}\right)+O\left(\frac{1}{n}\right)\right),$$

$$\mathbf{1}\{\mathcal{A}_\epsilon^1(n_0)\}\mathbb{E}(M(n+1)-M(n)\,|\,\mathcal{F}_n)\leq\mathbf{1}\{\mathcal{A}_\epsilon^1(n_0)\}\left(f_{\epsilon,1}^-\left(\frac{M(n)}{n}\right)+O\left(\frac{1}{n}\right)\right),$$

where

$$f_\epsilon^+(x)=m\left(1-\left(1-\frac{1}{2m+2k+\beta+\epsilon}x\right)^d\right),\qquad f_\epsilon^-(x)=m\left(1-\left(1-\frac{1}{2m+2k+\beta-\epsilon}x\right)^d\right).$$

The functions $f_\epsilon^\pm(x)$ are concave, $f_\epsilon^\pm(0)=0$, $(f_\epsilon^\pm)'(0)=\dfrac{dm}{2m+2k+\beta\pm\epsilon}$, and both $f_\epsilon^\pm(x)$ converge uniformly to $f(x)$ when $\epsilon\to0$. Also, $M(n+1)-M(n)<1+2m$. Therefore, due to Lemma 1, we get:

1. If $f'(0)<1$ than for any $\epsilon>0$, $f'(0)+\epsilon<1$,

$$\mathbb{P}\left(\forall n>n_0:n^{f'(0)-\epsilon}<Y(n)<n^{f'(0)+\epsilon}\right)\to1,\text{ as }n_0\to\infty.$$

2. If $f'(0)>1$ than for any $\epsilon>0$, $f'(0)-\epsilon>1$, almost surely,

$$\liminf_{n\to\infty}\frac{M(n)}{n}\geq x_\epsilon^+,\qquad\limsup_{n\to\infty}\frac{M(n)}{n}\leq x_\epsilon^-,$$

where x_ϵ^+ is a unique positive root of the equation $f_\epsilon^+(x)=x$, x_ϵ^- is a unique positive root of the equation $f_\epsilon^-(x)=x$.

Hence, by taking limit $\epsilon\to0$, we get the statements of Theorem 1.

The model studied in the article can be used, in particular, to describe long-range bonds in various nanostructured materials (for example, complex biological and polymeric structures).

The presented work was funded by a grant from the Russian Science Foundation (project No. 24-21-00247).

References:

1. Živkovic J., Tadic B. Nanonetworks: The graph theory framework for modeling nanoscale systems. *Nanoscale Systems: Mathematical Modeling, Theory and Applications*, 2013, vol. 2, pp. 30-48. DOI: 10.2478/nsmmt-2013-0003.
2. Vecchio D.A., Mahler S.H., Hammig M.D., Kotov N.A. Structural analysis of nanoscale network materials using graph theory. *ACS Nano*, 2021, vol. 15, issue 8, pp. 12847-12859. DOI: 10.1021/acsnano.1c04711.
3. Rana M.M., Nguyen D.D. Geometric graph learning to predict changes in binding free energy and protein thermodynamic stability upon mutation, *The Journal of Physical Chemistry Letters*, 2023, vol. 14, issue 49, pp. 10870-10879. DOI: 10.1021/acs.jpclett.3c02679.
4. Barabási A., Albert R. Emergence of scaling in random networks, *Science*, 1999, vol. 286, issue 5439, pp. 509-512. DOI: 10.1126/science.286.5439.509.
5. Móri T.F. On random trees, *Studia Scientiarum Mathematicarum Hungarica*, 2002, vol. 39, issue 1-2, pp. 143-155. DOI: 10.1556/SScMath.39.2002.1-2.9.
6. Móri T.F. The maximum degree of the Barabási-Albert random tree, *Combinatorics, Probability and Computing*, 2005, vol. 14, issue 3, pp. 339-348. DOI: 10.1017/S0963548304006133.
7. Haslegrave J., Jordan J. Preferential attachment with choice, *Random Structures and Algorithms*, 2016, vol. 48, issue 4, pp. 751-766. DOI: 10.1002/rsa.20616.
8. Krapivsky P.L, Redner S. Choice-driven phase transition in complex networks, *Journal of Statistical Mechanics: Theory and Experiment*, 2014, vol. 2014, art. no. P04021, 14 p. DOI: 10.1088/1742-5468/2014/04/P04021.
9. Malyshkin Y. Preferential attachment combined with the random number of choices, *Internet Mathematics*, 2018, vol. 1, issue 1, pp. 1-25. DOI: 10.24166/im.01.2018.
10. Malyshkin Y., Paquette E. The power of choice combined with preferential attachment, *Electronic Communications in Probability*, 2014, vol. 19, issue 44, pp. 1-13. DOI: 10.1214/ECP.v19-3461.
11. Haslegrave J., Jordan J., Yarrow M. Condensation in preferential attachment models with location-based choicem *Random Structures and Algorithms*, 2020, vol. 56, issue 3, pp. 775-795. DOI: 10.1002/rsa.20889.
12. Malyshkin Y. Sublinear preferential attachment combined with a growing number of choices, *Electronic Communications in Probability*, 2020, vol. 25, issue 87, pp. 1-12. DOI: 10.1214/20-ECP368.
13. Malyshkin Y., Paquette E. The power of choice over preferential attachment, *Latin American Journal of Probability and Mathematical Statistics*, 2015, vol. 12, issue 2, pp. 903-915.
14. Alves C., Ribeiro R., Sanchis R. Preferential attachment random graphs with edge-step functions, *Journal of Theoretical Probability*, 2021, vol. 34, issue 1, pp. 438-476. DOI: 10.1007/s10959-019-00959-0.
15. Chen H.F. Stochastic approximation and its applications, *Nonconvex Optimization and Its Applications*. New York, Springer, 2002, vol. 64, XV, 360 p. DOI: 10.1007/b101987.
16. Malyshkin Y. Preferential attachment with choice-based edge step. Available at: https://arxiv.org/abs/2309.16591 (accessed 01.08.2024). DOI: 10.48550/arXiv.2309.16591.

Библиографический список:

1. **Živkovic, J.** Nanonetworks: The graph theory framework for modeling nanoscale systems / J. Zivkovic, B. Tadic // Nanoscale Systems: Mathematical Modeling, Theory and Applications. – 2013. – V. 2. – P. 30-48. DOI: 10.2478/nsmmt-2013-0003.
2. **Vecchio, D.A.** Structural analysis of nanoscale network materials using graph theory / D.A. Vecchio, S.H. Mahler, M.D. Hammig, N.A. Kotov // ACS Nano. – 2021. – V. 15 – I. 8. – P. 12847-12859. DOI: 10.1021/acsnano.1c04711.
3. **Rana, M.M.** Geometric graph learning to predict changes in binding free energy and protein thermodynamic stability upon mutation / M.M. Rana, D.D. Nguyen // The Journal of Physical Chemistry Letters. – 2023. – V. 14 – I. 49. – P. 10870-10879. DOI: 10.1021/acs.jpclett.3c02679.
4. **Barabási, A.** Emergence of scaling in random networks / A. Barabási, R. Albert // Science. – 1999. – V. 286. – I. 5439. – P. 509-512. DOI: 10.1126/science.286.5439.509.
5. **Móri, T.F.** On random trees / T.F. Móri // Studia Scientiarum Mathematicarum Hungarica. – 2002. – V. 39. – I. 1-2. – P. 143-155. DOI: 10.1556/SScMath.39.2002.1-2.9
6. **Móri, T.F.** The maximum degree of the Barabási-Albert random tree / T.F. Móri // Combinatorics, Probability and Computing. – 2005. – V. 14. – I. 3. – P. 339-348. DOI: 10.1017/S0963548304006133.
7. **Haslegrave, J.** Preferential attachment with choice / J. Haslegrave, J. Jordan // Random Structures and Algorithms. – 2016. – V. 48. – I. 4. – P. 751-766. DOI: 10.1002/rsa.20616.
8. **Krapivsky, P.L.** Choice-driven phase transition in complex networks / P.L. Krapivsky, S. Redner // Journal of

Statistical Mechanics: Theory and Experiment. – 2014. – V. 2014. – Art. № P04021. – 14 p. DOI: 10.1088/1742-5468/2014/04/P04021.
9. **Malyshkin, Y.** Preferential attachment combined with the random number of choices / Y. Malyshkin // Internet Mathematics. – 2018. – V. 1. – I. 1. – P. 1-25. DOI: 10.24166/im.01.2018.
10. **Malyshkin, Y.** The power of choice combined with preferential attachment / Y. Malyshkin, E. Paquette // Electronic Communications in Probability. – 2014. – V. 19. – I. 44. – P. 1-13. DOI: 10.1214/ECP.v19-3461.
11. **Haslegrave, J.** Condensation in preferential attachment models with location-based choice / J. Haslegrave, J. Jordan, M. Yarrow // Random Structures and Algorithms. – 2020. – V. 56. – I. 3. – P. 775-795. DOI: 10.1002/rsa.20889.
12. **Malyshkin, Y.** Sublinear preferential attachment combined with a growing number of choices / Y. Malyshkin // Electronic Communications in Probability. – 2020. – V. 25. – I. 87. – P. 1-12. DOI: 10.1214/20-ECP368.
13. **Malyshkin, Y.** The power of choice over preferential attachment / Y. Malyshkin, E Paquette. // Latin American Journal of Probability and Mathematical Statistics. – 2015. – V. 12. – I. 2. – P. 903-915.
14. **Alves, C.** Preferential Attachment Random Graphs with Edge-Step Functions / C. Alves, R. Ribeiro, R. Sanchis // Journal of Theoretical Probability. – 2021. – V. 34. – I. 1. – P. 438-476. DOI: 10.1007/s10959-019-00959-0.
15. **Chen, H.F.** Stochastic approximation and its applications / H.F. Chen // In book series: Nonconvex Optimization and Its Applications. – New York: Springer, 2002. – V. 64. – XV, 360 p. DOI: 10.1007/b101987.
16. **Malyshkin, Y.** Preferential attachment with choice-based edge step / Y. Malyshkin. – Режим доступа: https://arxiv.org/abs/2309.16591. – 01.08.2024. DOI: 10.48550/arXiv.2309.16591.

УДК 621.892 *Краткое сообщение*

Геометрическое предпочтительное присоединение с добавлением ребер, основанном на выборе вершины

Ю.А. Малышкин
ФГБОУ ВО «Тверской государственный университет»
170002, Россия, Тверь, Садовый пер., 35

DOI: 10.26456/pcascnn/2024.16.381

Аннотация: Мы изучаем асимптотическое поведение максимальной степени вершины в геометрической модели графа с шагом добавления ребра на основе выбора вершины. Геометрические графы являются естественными моделями, которые описывают наномасштабные системы, в то время как предпочтительное присоединение обеспечивает хорошее описание сложных сетей, в частности, различных нейронных сетей. Модель представляет собой рекурсивно построенную последовательность графов. Мы начинаем с исходного графа на одной вершине и на каждом шаге добавляем новую вершину и проводим несколько ребер. Каждой вершине соответствует параметр, характеризующий ее местоположение. Шаг рекурсии состоит из двух частей. Сначала мы вводим новую вершину и проводим ребра от нее до достаточно близких вершин. Затем мы проводим ребро между вершинами с помощью предпочтительного присоединения с выбором. В работе доказывается, что в зависимости от параметров модели максимальная степень может демонстрировать сублинейное (аналогично стандартному предпочтительному присоединению) и линейное (что представляет эффект концентрации) поведение.

Ключевые слова: геометрические графы, сложные сети, случайные графы, предпочтительное присоединение, выбор вершины.

Малышкин Юрий Андреевич – к.ф.-м.н.., доцент кафедры информационных технологий ФГБОУ ВО «Тверской государственный университет»

Yury A. Malyshkin – Ph. D., Docent of Applied Mathematics and Cybernetics Department, Tver State University

Поступила в редакцию/received: 01.07.2024; после рецензирования/revised: 02.08.2024; принята/accepted 09.08.2024.