

А. А. Печников (Петрозаводск, ИПМИ КарНЦ РАН). **Метрические характеристики динамического графа цитирования Math-Net.Ru.**

УДК 004.9:519.178

Резюме: Исследованы характеристики динамического графа цитирования Math-Net.Ru, построенного по данным за последние пятьдесят лет. Показано, что средняя длина пути и диаметр графа практически стабильны во времени, имея значения около $5,4(\pm 1)$ и $17,5(\pm 3)$ соответственно, а доля вершин максимальной компоненты связности в общем числе вершин после начального двадцатилетнего периода накопления имеет тенденцию к стабилизации в интервале от 30 до 40%. Предложено содержательное объяснение стабильности этих характеристик.

Ключевые слова: динамический граф, граф цитирования, характеристики графа.

Графом цитирования называется ориентированный граф, вершины которого соответствуют авторам научных публикаций, а дуга (i, j) существует тогда и только тогда, когда автор i хотя бы одну ссылку, цитирующую автора j . Вес дуги равен количеству ссылок во всех публикациях автора i на публикации автора j . По определению в графе цитирования отсутствуют изолированные вершины. Граф цитирования не содержит петель, поскольку самоцитирование не рассматривается в данной работе.

Зафиксируем год T , и пусть на этот год имеется граф цитирования $G_N(V_T, E_T, W_T)$. Выберем шаг Δt — натуральное число, и рассмотрим граф цитирования $G_{N-1}(V_{T-\Delta t}, E_{T-\Delta t}, W_{T-\Delta t})$ на год $T - \Delta t$. Множество дуг $E_{T-\Delta t}$ получается из E_T путем удаления всех дуг, соответствующих ссылкам, сделанным за период $(T - \Delta t, T]$, а множество $V_{T-\Delta t}$ — удалением всех изолированных вершин из V_T , ставших таковыми после удаления дуг. Значения элементов множества $W_{T-\Delta t}$ равны кратности дуг в $E_{T-\Delta t}$. Для графа G_{N-1} можно построить граф G_{N-2} , и т. д., получив последовательность $G_{N-k}, \dots, G_{N-1}, G_N$, ($k < N$), называемую динамическим графом (сетью) [1]. Одной из задач для динамических графов является задача обнаружения наследственности, когда при переходе от предыдущего к последующему графу сохраняются некоторые характеристики и устанавливаются операции перехода, сохраняющие их.

По данным Math-Net.Ru [2] была построен динамический граф, состоящий из 50 графов цитирования с 1970 по 2019 год. Для примера: количество вершин в графе G_{1970} равно 2583, дуг 9537, кратных дуг — 18065, а в графе G_{2019} — 52721 вершины, 388606 дуг и 862340 кратных дуг. Количество вершин во времени аппроксимируется линейной функцией, а количество дуг и кратных дуг — полиномиальными функциями третьего и четвертого порядка соответственно. В графе на протяжении почти 50 лет практически сохраняются такие характеристики, как средняя длина пути (около $5,4 \pm 1$) и диаметр (около $17,5 \pm 3$). Такая вторичная характеристика, как доля вершин максимальной компоненты сильной связности (МКСС) после начального двадцатилетнего периода накопления имеет тенденцию к стабилизации в интервале от 30 до 40%.

Граф не соответствует модели «малого мира» [3] и требуются другие модели для объяснения его динамики во времени. Одним из объяснений может быть следующее:

- 1) новая статья текущего года с большой вероятностью ссылается на статьи, входящие в МКСС, учитывая их накопленный авторитет;
- 2) некоторое время ссылок на новую статью не будет, и она будет «висячей», таким образом, увеличивая диаметр и среднюю длину пути,
- 3) в то же время для статей предыдущих лет увеличивается вероятность появления ссылок на них, что ведет к уменьшению диаметра и средней длины пути.

Работа выполнена при поддержке РФФИ, проект 18-07-00628-а.

СПИСОК ЛИТЕРАТУРЫ

1. *Boccaletti S., Latora V., Moreno Y., Chavez M., Hwang D. U.* Complex networks: Structure and function. — *Phys Rep*, 2006, v. 424, p. 175–308.
2. *Chebukov D., Izaak A., Misyrina O., Pupyrev Yu., Zhizhchenko A.* Math-Net.Ru as a digital archive of the Russian mathematical knowledge from the XIX century to today. — *Lecture Notes in Computer Science*, 2013, v. 7961, p. 344–348.
3. *Humphries M. D., Gurney K.* Network ‘Small-World-Ness’: A Quantitative Method for Determining Canonical Network Equivalence. — *PLoS ONE*, 2008, v. 3(4):e0002051. doi:10.1371/journal.pone.0002051.

УДК УДК 004.9:519.178

***Pechnikov A. A.* (Petrozavodsk, IAMR KarRC RAS). Metric characteristics of the dynamic citation graph of Math-Net.Ru**

Abstract: This paper investigates the characteristics of the dynamic citation graph of Math-Net.Ru constructed based on the data collected for the last fifty years. It is shown that the average path length and diameter of the graph are almost stable in time — about $5.4(\pm 1)$ and $17.5(\pm 3)$ respectively — and the proportion of vertices of the maximum connected component in the total number of vertices after the initial twenty years of accumulation tends to stabilize in the interval from 30 to 40%. An in-depth explanation of the stability of these characteristics is proposed.

Keywords: dynamic graph, citation graph, characteristics of graph.