УДК 510.52

О ГЕНЕРИЧЕСКОЙ СЛОЖНОСТИ ПРОБЛЕМЫ РАСПОЗНАВАНИЯ ГАМИЛЬТОНОВЫХ ПУТЕЙ[1]

А. Н. Рыбалов

Институт математики им. С. Л. Соболева СО РАН, г. Омск, Россия

Изучается генерическая сложность проблемы распознавания гамильтоновых путей в конечных графах. Путь в графе называется гамильтоновым, если он проходит через все вершины ровно по одному разу. Доказывается, что при условии $\mathrm{P} \neq \mathrm{NP}$ и $\mathrm{P} = \mathrm{BPP}$ для этой проблемы не существует полиномиального сильно генерического алгоритма.

Ключевые слова: *генерическая сложность, гамильтонов путь.*

DOI 10.17223/20710410/53/8

THE GENERAL COMPLEXITY OF THE PROBLEM TO RECOGNIZE HAMILTONIAN PATHS

A. N. Rybalov

Sobolev Institute of Mathematics, Omsk, Russia

E-mail: alexander.rybalov@gmail.com

Generic-case approach to algorithmic problems has been offered by A. Miasnikov, V. Kapovich, P. Schupp, and V. Shpilrain in 2003. This approach studies an algorithm behavior on typical (almost all) inputs and ignores the rest of inputs. In this paper, we study the generic complexity of the problem of recognition of Hamiltonian paths in finite graphs. A path in graph is called Hamiltonian if it passes through all vertices exactly once. We prove that under the conditions $\mathrm{P} \neq \mathrm{NP}$ and $\mathrm{P} = \mathrm{BPP}$ for this problem there is no polynomial strongly generic algorithm. A strongly generic algorithm solves a problem not on the whole set of inputs, but on a subset, the sequence of frequencies of which exponentially quickly converges to 1 with increasing size. To prove the theorem, we use the method of generic amplification, which allows to construct generically hard problems from the problems hard in the classical sense. The main component of this method is the cloning technique, which combines the inputs of a problem together into sufficiently large sets of equivalent inputs. Equivalence is understood in the sense that the problem is solved similarly for them.

Keywords: *generic complexity, Hamiltonian path.*

Введение

Гамильтонов путь в графе — это путь, проходящий через все вершины графа ровно по одному разу. Проблема распознавания гамильтоновых путей в конечных графах является классической комбинаторной проблемой, изучаемой многие десятилетия. Она содержится в знаменитом списке из двадцати одной NP-полной проблемы [1]. При

[1]Работа поддержана грантом РНФ № 19-11-00209.

условии $\mathrm{P} \neq \mathrm{NP}$ не существует полиномиального алгоритма для решения этой задачи. Здесь P — это класс алгоритмических проблем, распознаваемых за полиномиально ограниченное время на детерминированных машинах Тьюринга, а класс NP состоит из проблем, распознаваемых за полиномиальное время на недетерминированных машинах Тьюринга. Эквивалентное определение класса NP — это класс проблем, решения которых можно проверить за полиномиальное время на детерминированных машинах Тьюринга. Большинство исследователей считает, что имеет место неравенство $\mathrm{P} \neq \mathrm{NP}$, однако это до сих пор не доказано. При условии совпадения классов P и BPP (класс проблем, решаемых за полиномиальное время вероятностными алгоритмами) для проблемы распознавания гамильтоновых путей не существует и полиномиальных вероятностных алгоритмов. Имеются серьёзные доводы в пользу равенства $\mathrm{P} = \mathrm{BPP}$. В частности, доказано [2], что это равенство следует из весьма правдоподобных гипотез о вычислительной сложности некоторых трудных проблем.

В [3] предложена теория генерической вычислимости и сложности вычислений. В рамках этого подхода алгоритмическая проблема рассматривается не на всём множестве входов, а на некотором подмножестве «почти всех» входов. Такие входы образуют так называемое генерическое множество. Понятие «почти все» формализуется введением естественной меры на множестве входных данных. С точки зрения практики алгоритмы, решающие быстро проблему на генерическом множестве, так же хороши, как и быстрые алгоритмы для всех входов. Отметим, что похожий подход для изучения проблем оптимизации был предложен ранее в [4].

Данная работа посвящена изучению генерической сложности распознавания гамильтоновых путей. Доказывается, что при условии $\mathrm{P} \neq \mathrm{NP}$ и $\mathrm{P} = \mathrm{BPP}$ для этой проблемы не существует полиномиального сильно генерического алгоритма. Сильно генерический алгоритм решает проблему не на всём множестве входов, а на подмножестве, последовательность частот которого при увеличении размера экспоненциально быстро сходится к 1.

1. Предварительные сведения

Пусть I — некоторое множество входов, I_n — подмножество входов размера n. Для подмножества $S \subseteq I$ определим последовательность

$$\rho_n(S) = \frac{|S_n|}{|I_n|},\ n = 1, 2, 3, \ldots,$$

где $S_n = S \cap I_n$ — множество входов из S размера n. Заметим, что $\rho_n(S)$ — это вероятность попасть в S при случайной и равновероятной генерации входов из I_n. *Асимптотической плотностью* S назовём предел

$$\rho(S) = \overline{\lim_{n\to\infty}} \rho_n(S).$$

Верхний предел здесь нужен для того, что часто при кодировании входных данных не для каждого n существуют коды размера n. Множество S называется *пренебрежимым*, если $\rho(S) = 0$, и *сильно пренебрежимым*, если последовательность $\rho_n(S)$ экспоненциально быстро сходится к 0, т. е. существуют константы σ, $0 < \sigma < 1$, и $C > 0$, такие, что для любого n имеет место $\rho_n(S) < C\sigma^n$.

Алгоритм $\mathcal{A}$ с множеством входов I и множеством выходов $J \cup \{?\}$ $(? \notin J)$ называется (*сильно*) *генерическим*, если:

1) $\mathcal{A}$ останавливается на всех входах из I;

2) множество $\{x \in I : \mathcal{A}(x) = ?\}$ является (сильно) пренебрежимым.

Генерический алгоритм $\mathcal{A}$ вычисляет функцию $f : I \to J$, если для всех $x \in I$ выполнено

$$(\mathcal{A}(x) = y \in J) \Rightarrow (f(x) = y).$$

Ситуация $\mathcal{A}(x) = ?$ означает, что $\mathcal{A}$ не может вычислить функцию f на аргументе x. Но условие 2 гарантирует, что $\mathcal{A}$ корректно вычисляет f на почти всех входах (входах из генерического множества). Различие между генерически разрешимыми проблемами и сильно генерически разрешимыми проблемами поясняется в работе [5].

2. Гамильтоновы пути

Здесь и далее будем рассматривать неориентированные графы без петель и кратных рёбер. Пусть имеется граф G с множеством вершин $\{v_1, \ldots, v_n\}$. Путь с началом в вершине v_1 и концом в вершине v_n, проходящий через каждую вершину ровно по одному разу, называется *гамильтоновым*. *Проблема распознавания гамильтоновых путей* состоит в следующем. Задан произвольный граф G. Определить, существует ли в графе G гамильтонов путь.

Обычно в определении гамильтонова пути не накладывается условие на начало и конец. Однако, как показывает следующая лемма, это условие не делает проблему более простой, то есть проблема распознавания таких гамильтоновых путей тоже неразрешима за полиномиальное время при условии $\mathrm{P} \neq \mathrm{NP}$.

Лемма 1. Если для проблемы распознавания гамильтоновых путей с выделенными началом и концом существует полиномиальный алгоритм, то существует полиномиальный алгоритм и для проблемы распознавания гамильтоновых путей без выделенных начала и конца.

Доказательство. Допустим, что существует полиномиальный алгоритм $\mathcal{A}$ для проблемы распознавания наличия гамильтоновых путей с выделенными началом и концом. Построим полиномиальный алгоритм $\mathcal{B}$ для проблемы распознавания наличия гамильтоновых путей без выделенных начала и конца.

Алгоритм $\mathcal{B}$ работает на графе G с вершинами $v_1, \ldots, v_n$ следующим образом. Для всевозможных $i, j \leqslant n$, таких, что $i \neq j$, строится граф $G(i, j)$, который получается из графа G добавлением двух новых вершин u, w и двух новых рёбер (u, v_i) и (v_j, w). Вершина u объявляется первой, а вершина w — последней вершинами нового графа $G(i, j)$. Затем запускается алгоритм $\mathcal{A}$ на графе $G(i, j)$. Если хотя бы на одном графе $G(i, j)$ алгоритм $\mathcal{A}$ определил, что существует гамильтонов путь с началом в u и концом в w, то существует и гамильтонов путь в графе G. Если же для всех графов $G(i, j)$ таких путей нет, то нет гамильтонова пути и в графе G.

Этот алгоритм является полиномиальным, так как алгоритм $\mathcal{A}$ является полиномиальным и число графов $G(i, j)$ равно $n^2 - n$. ■

Пусть G_1 и G_2 — два графа с непересекающимися множествами вершин и v_1 — некоторая вершина графа G_1, а v_2 — некоторая вершина графа G_2. Обозначим через $G_1 \cup G_2 \cup (v_1, v_2)$ граф, являющийся объединением графов G_1 и G_2 с добавленным ребром (v_1, v_2).

Лемма 2. Пусть G_1 и G_2 — два графа с непересекающимися множествами вершин $\{v_1, \ldots, v_k\}$ и $\{w_1, \ldots, w_m\}$ соответственно. Пусть в графе G_2 существует гамильтонов путь. Тогда в графе G_1 существует гамильтонов путь тогда и только тогда, когда в графе $G_1 \cup G_2 \cup (v_k, w_1)$ существует гамильтонов путь.

Доказательство. Пусть $\pi = w_1 w'_2 \dots w'_{m-1} w_m$ — гамильтонов путь в графе G_2, $\tau = v_1 v'_2 \dots v'_{k-1} v_k$ — гамильтонов путь в графе G_1. Тогда легко видеть, что в графе $G_1 \cup G_2 \cup (v_k, w_1)$ гамильтоновым путём является путь

$$\psi = v_1 v'_2 \dots v'_{k-1} v_k w_1 w'_2 \dots w'_{m-1} w_m.$$

Обратно, пусть в графе $G_1 \cup G_2 \cup (v_k, w_1)$ есть гамильтонов путь $\tau = v_1 u_1 \dots u_{k+m-2} w_k$, где $u_i \in \{v_2, \dots, v_k, w_1, \dots, w_{k-1}\}$, $i = 1, \dots, k+m-2$. Тогда этот путь обязательно проходит через ребро (v_k, w_1), так как это единственное ребро, соединяющее графы G_1 и G_2. Поэтому этот путь имеет вид

$$\tau = v_1 v'_2 \dots v'_{k-1} v_k w_1 w'_2 \dots w'_{m-1} w_m,$$

где $v'_i \in \{v_2, \dots, v_{k-1}\}$, а $w'_i \in \{w_2, \dots, w_{m-1}\}$. Отсюда следует, что путь $\pi = v_1 v'_2 \dots v'_{k-1} v_k$ является гамильтоновым для графа G_1. ■

Лемма 3. Пусть $\mathcal{G}_n$ — множество всех графов, а $\mathcal{H}_n$ — множество графов, имеющих гамильтонов путь, с вершинами $v_1, \dots, v_n$. Тогда для достаточно больших n имеет место оценка

$$\frac{|\mathcal{H}_n|}{|\mathcal{G}_n|} > \frac{0{,}99}{n^2}.$$

Доказательство. Известно [6, 7], что почти во всех графах при достаточно большом n существует гамильтонов путь с началом в некоторой вершине v_i и концом в некоторой вершине v_j, $i \neq j$, $i, j \leqslant n$. Обозначим множество таких графов через $\mathcal{H}'_n$. Имеем

$$\frac{|\mathcal{H}'_n|}{|\mathcal{G}_n|} > 0{,}99. \tag{1}$$

Нас интересует множество графов $\mathcal{H}_n$, в которых гамильтонов путь начинается в вершине v_1 и заканчивается в вершине v_n. Обозначим через $\mathcal{H}_n(i, j)$ множество графов с гамильтоновым путём, начинающимся в вершине v_i и заканчивающимся в вершине v_j, $i \neq j$. Таким образом, $\mathcal{H}_n = \mathcal{H}_n(1, n)$. Заметим, что

$$\mathcal{H}'_n = \bigcup_{i,j \leqslant n,\ i \neq j} \mathcal{H}_n(i, j).$$

Отсюда

$$|\mathcal{H}'_n| \leqslant \sum_{i,j \leqslant n,\ i \neq j} |\mathcal{H}_n(i, j)|. \tag{2}$$

Для любых i, j, $i \neq j$, имеет место $|\mathcal{H}_n(i, j)| = |\mathcal{H}_n(1, n)|$. Действительно, перестановка на множестве индексов, которая переводит i в 1, j в n, а остальные индексы оставляет на месте, задаёт биекцию между множествами $\mathcal{H}_n(i, j)$ и $\mathcal{H}_n(1, n)$. Поэтому неравенство (2) можно переписать как

$$|\mathcal{H}'_n| \leqslant (n^2 - n)|\mathcal{H}_n(1, n)| = (n^2 - n)|\mathcal{H}_n|.$$

С учётом (1) получаем

$$\frac{|\mathcal{H}_n|}{|\mathcal{G}_n|} \geqslant \frac{0{,}99}{n^2 - n} > \frac{0{,}99}{n^2}.$$

Лемма 3 доказана. ■

3. Основной результат

Для изучения генерической сложности проблемы распознавания гамильтоновых путей будем использовать представление графов с помощью матриц смежности. Так как графы неориентированные, для кодирования графа с n вершинами достаточно верхней части такой матрицы, состоящей из $n(n-1)/2$ бит. Таким образом, будем считать, что размер графа с n вершинами равен $n(n-1)/2$.

Теорема 1. Если существует сильно генерический полиномиальный алгоритм, решающий проблему распознавания гамильтоновых путей, то существует вероятностный полиномиальный алгоритм, решающий эту проблему на всём множестве входов.

Доказательство. Допустим, что существует сильно генерический полиномиальный алгоритм $\mathcal{A}$, решающий проблему распознавания гамильтоновых путей. Построим вероятностный полиномиальный алгоритм $\mathcal{B}$, разрешающий эту проблему на всём множестве входов. Пусть имеется граф G с n вершинами $v_1, \ldots, v_n$. Он имеет размер $n(n-1)/2$. Алгоритм $\mathcal{B}$ будет работать на графе G следующим образом:

1. Генерируется случайный граф H, имеющий гамильтонов путь, с n^2-n вершинами $v_{n+1}, \ldots, v_{n^2}$. Это делается за $n(n^2-n)^2$ раундов:
   а) генерируется случайный граф H с n^2-n вершинами $v_{n+1}, \ldots, v_{n^2}$;
   б) с помощью алгоритма $\mathcal{A}$ проверяется, есть ли в графе H гамильтонов путь;
   в) если $\mathcal{A}(H) =?$, то возвращаемся на шаг а;
   г) если $\mathcal{A}(H) = 0$, то возвращаемся на шаг а;
   д) если $\mathcal{A}(H) = 1$, то генерация заканчивается.
2. Если генерация прошла успешно, переходим на следующий шаг, иначе выдаём ответ «НЕТ».
3. Запускается алгоритм $\mathcal{A}$ на графе $G' = G \cup H \cup (v_n, v_{n+1})$.
4. Если $\mathcal{A}(G') = 1$, то, по лемме 2, в графе G есть гамильтонов путь. Выдаём ответ «ДА».
5. Если $\mathcal{A}(G') = 0$, то, по лемме 2, в графе G нет гамильтонова пути. Выдаём ответ «НЕТ».
6. Если $\mathcal{A}(G') =?$, то выдаём ответ «НЕТ».

Заметим, что полиномиальный вероятностный алгоритм $\mathcal{B}$ выдаёт правильный ответ на шагах 4 и 5, а на шагах 2 и 6 может выдать неправильный ответ. Нужно доказать, что вероятность того, что ответ выдаётся на шагах 2 и 6, меньше $1/2$.

Граф H имеет n^2-n вершин, то есть его размер равен $m = (n^2-n)(n^2-n-1)/2$. Так как множество $\{G \in \mathcal{G} : \mathcal{A}(G) \neq ?\}$ сильно пренебрежимое, то существует константа $\alpha > 0$, такая, что

$$\frac{|\{H \in \mathcal{G} : \mathcal{A}(H) \neq ?\}_m|}{|\mathcal{G}_m|} < \frac{1}{2^{\alpha m}} = \frac{1}{2^{\alpha(n^2-n)(n^2-n-1)/2}}$$

для любого n.

Ответ будет выдан на шаге 2 в том случае, если за все $n(n^2-n)^2$ раундов генерации не был получен граф H с гамильтоновым путём. Эта вероятность, по лемме 3, для достаточно больших n не превосходит

$$\left(1 - \frac{0{,}99}{(n^2-n)^2} + \frac{1}{2^{\alpha(n^2-n)(n^2-n-1)/2}}\right)^{n(n^2-n)^2} < \left(1 - \frac{0{,}98}{(n^2-n)^2}\right)^{n(n^2-n)^2} < \frac{1}{\mathrm{e}^{0{,}98n}} < \frac{1}{4}.$$

Оценим вероятность выдать ответ на шаге 6. Граф $G \cup H \cup (v_n, v_{n+1})$ имеет n^2 вершин, то есть его размер равен $m = (n^4 - n^2)/2$. Вероятность того, что для случайного графа $G \cup H \cup (v_n, v_{n+1})$ имеет место $\mathcal{A}(G \cup H \cup (v_n, v_{n+1})) =?$, не больше

$$\frac{|\{G \in \mathcal{G} : \mathcal{A}(G) \neq ?\}_m|}{|\{G \cup H \cup (v_n, v_{n+1}) : H \in \mathcal{G}\}_m|} = \frac{|\{G \in \mathcal{G} : \mathcal{A}(G) \neq ?\}_m|}{|\mathcal{G}_m|} \frac{|\mathcal{G}_m|}{|\{G \cup H \cup (v_n, v_{n+1}) : H \in \mathcal{G}\}_m|}.$$

С другой стороны, так как граф H имеет $n^2 - n$ вершин, то

$$|\{G \cup H \cup (v_n, v_{n+1}) : H \in \mathcal{G}\}_m| = |\{H : H \in \mathcal{G}\}_{((n^2-n)^2-(n^2-n))/2}| = 2^{(n^4-2n^3+n)/2}.$$

Отсюда

$$\frac{|\mathcal{G}_m|}{|\{G \cup H \cup (v_n, v_{n+1}) : H \in \mathcal{G}\}_m|} = \frac{2^{(n^4-n^2)/2}}{2^{(n^4-2n^3+n)/2}} = 2^{(2n^3-n^2+n)/2}.$$

Поэтому искомая вероятность не больше

$$\frac{2^{(2n^3-n^2+n)/2}}{2^{\alpha(n^4-n^2)/2}} < \frac{1}{4}$$

при больших n.

Итого, получаем, что вероятность выдать ответ на шагах 2 и 6 не превосходит $1/4 + 1/4 = 1/2$. ■

Напомним, что класс BPP состоит из проблем, разрешимых за полиномиальное время на вероятностных машинах Тьюринга. Одной из важных гипотез в теории сложности вычислений является гипотеза о совпадении классов P и BPP. Из неё следует, что любой полиномиальный вероятностный алгоритм $\mathcal{A}$ можно эффективно дерандомизировать, то есть построить полиномиальный алгоритм $\mathcal{B}$, не использующий генератор случайных чисел и решающий ту же проблему, что и алгоритм $\mathcal{A}$. В работе [2] доказано, что равенство P = BPP следует из весьма правдоподобных гипотез о вычислительной сложности некоторых трудных проблем.

Теорема 2. Если P $\neq$ NP и P=BPP, то не существует сильно генерического полиномиального алгоритма для решения проблемы распознавания гамильтоновых путей.

Доказательство. Пусть существует сильно генерический алгоритм, решающий проблему распознавания гамильтоновых путей. Тогда по теореме 1 существует вероятностный полиномиальный алгоритм, решающий её на всём множестве входов, т. е. эта проблема лежит в классе BPP. Так как P = BPP, она лежит и в классе P, откуда, по лемме 1, P $\neq$ NP, что противоречит посылке теоремы. ■

Автор выражает благодарность рецензенту за полезные замечания и предложения по улучшению текста статьи.

ЛИТЕРАТУРА

1. *Karp R.* Reducibility among combinatorial problems // R. E. Miller and J. W. Thather (eds.). Complexity of Computer Computations. The IBM Research Symposia Ser., 1972. P. 85–103.
2. *Impagliazzo R. and Wigderson A.* P = BPP unless E has subexponential circuits: Derandomizing the XOR Lemma // Proc. 29th STOC. El Paso: ACM, 1997. P. 220–229.
3. *Kapovich I., Miasnikov A., Schupp P., and Shpilrain V.* Generic-case complexity, decision problems in group theory and random walks // J. Algebra. 2003. V. 264. No. 2. P. 665–694.
4. *Гимади Э.Х., Глебов Н.И., Перепелица В.А.* Алгоритмы с оценками для задач дискретной оптимизации // Проблемы кибернетики. 1975. Т. 31. С. 35–42.

5. *Рыбалов А.* О генерической сложности проблемы общезначимости булевых формул // Прикладная дискретная математика. 2016. № 2(32). С. 119–126.
6. *Перепелица В.А.* О двух задачах из теории графов // Докл. АН СССР. 1970. Т. 194. № 6. С. 1269–1272.
7. *Posa L.* Hamiltonian circuits in random graphs // Discrete Math. 1976. V. 14. P. 359–364.

REFERENCES

1. *Karp R.* Reducibility among combinatorial problems. R. E. Miller and J. W. Thather (eds.), Complexity of Computer Computations. The IBM Research Symposia Ser., 1972, pp. 85–103.
2. *Impagliazzo R. and Wigderson A.* P = BPP unless E has subexponential circuits: Derandomizing the XOR Lemma. Proc. 29th STOC, El Paso, ACM, 1997, pp. 220–229.
3. *Kapovich I., Miasnikov A., Schupp P., and Shpilrain V.* Generic-case complexity, decision problems in group theory and random walks. J. Algebra, 2003, vol. 264, no. 2, pp. 665–694.
4. *Gimadi E.H., Glebov N.I., and Perepelitsa V.A.* Algoritmy s ocenkami dlya zadach diskretnoi optimizacii [Algorithms with bounds for problems of discrete optimization]. Problemy Kibernetiki, 1975, vol. 31, pp. 35–42. (in Russian)
5. *Rybalov A.* O genericheskoy slozhnosti problemy obshcheznachimosti bulevykh formul [On generic complexity of the validity problem for Boolean formulas]. Prikladnaya Diskretnaya Matematika, 2016, no. 2(32), pp. 119–126. (in Russian)
6. *Perepelitsa V.A.* O dvuh zadachah iz teorii grafov [On two problems from graph theory]. Doklady AN SSSR, 1970, vol. 194. no. 6, pp. 1269–1272. (in Russian)
7. *Posa L.* Hamiltonian circuits in random graphs. Discrete Math., 1976, vol. 14, pp. 359–364.