О задаче приближенного нахождения максимальной двудольной клики

Н.Н. Кузюрин
Интитут системного программирования РАН,
109004, Москва, ул. А. Солженицына, 25

Аннотация. Задача о нахождении большой "спрятанной" клики в случайном графе и ее аналог для двудольных графов являются объектами рассмотрения в данной заметке.

Ключевые слова: случайный граф; большая спрятанная клика; сложность нахождения

DOI: 10.15514/ISPRAS-2017-29(3)-12

Для цитирования: Кузюрин Н.Н. О задаче приближенного нахождения максимальной двудольной клики. Труды ИСП РАН, том 29, вып. 3, 2017 г., стр. 225-232. DOI: 10.15514/ISPRAS-2017-29(3)-12

1. Введение

Изучение свойств случайных дискретных структур является важным направлением дискретной математики, в последние годы привлекающим внимание все большего числа исследователей во всем мире. В целом, роль вероятностных методов в современном развитии дискретной математики трудно переоценить: эти методы используются как при изучении различных свойств случайных структур (графов, гиперграфов и т.д.), так и при доказательствах существования комбинаторных объектов с заданными свойствами и построении эффективных алгоритмов. Заметную роль здесь играют несколько трудных задач, которые привлекли внимание специалистов и не поддаются решению. Одной таких задач является задача о нахождении большой "спрятанной" клики в случайном графе. Именно эта задача и ее аналог для двудольных графов являются объектами рассмотрения в данной заметке.

2. Обзор известных результатов

Прежде чем говорить о сложности задач для случайных графов рассмотрим, что известно о трудности нахождения их приближенных решений при анализе по худшему случаю.

Определение. *Кликой в графе G называется множество вершин, любые две из которых соединены ребром. Максимально возможное число вершин в клике G обозначается $w(G)$ и называется размером максимальной клики.*

Сформулируем сейчас аналог задачи о максимальной клике для случая двудольных графов. Напомним, что граф $G = (V, E)$ называется двудольным, если его множество вершин можно разбить на два непустых подмножества V_1 и V_2 так, что $V = V_1 \cup V_2$ и в графе нет ребра, оба конца которого принадлежат одному из множеств V_1 или V_2.

Определение. *Двудольной кликой в двудольном графе $G = (V_1, V_2, E)$ называется полный двудольный подграф (U, W) графа $G, U \subseteq V_1, W \subseteq V_2$*

Известно несколько вариантов задачи о двудольной клике.

Сбалансированная двудольная клика. В этом случае $|V_1| = |V_2|$ и $|U| = |W|$. Известно, что задача нахождения максимальной двудольной сбалансированной клики (т.е. максимизации $|U| = |W|$) NP-полна [6].

Другой вариант задачи – это двудольная клика с максимальным числом ребер, т.е требуется максимизировать $|U| \cdot |W|$. Эта задача, как показано сравнительно недавно, также NP-полна [3].

Известно, что задача о максимальной клике в произвольном графе NP-полна [6]. В результате длительных исследований удалось доказать, что задача о максимальной клике очень плохо аппроксимируется. Напомним, что мультипликативной ошибкой алгоритма A называется максимум по всем входам данной длины отношения стоимости решения, найденного алгоритмом A, к стоимости оптимального решения. Наилучший известный полиномиальный алгоритм для нахождения максимальной клики гарантирует мультипликативную ошибку не более $O(n(\log\log n)^2/(\log n)^3)$ [4]. Отметим, что аппроксимация с ошибкой n тривиальна, так, что полученный результат не намного улучшает тривиальную оценку. Более того, Хостад [10] показал, что для задачи о максимальной клике не существует полиномиального приближенного алгоритма, имеющего ошибку менее $n^{1-\delta}$, для любого фиксированного $\delta > 0$ (в предположении $RP \neq NP$).

Известно, однако, что для ряда задач на графах их аналоги для двудольных графов решаются значительно проще. Возможно именно этим объясняется тот факт, что результаты о трудности аппроксимации для задачи СБАЛАНСИРОВАННАЯ ДВУДОЛЬНАЯ КЛИКА гораздо слабее результатов для задачи о клике в произвольном графе. Приведем сейчас известные результаты о трудности аппроксимации задачи о сбалансированной двудольной клике.

1. В работе [1] доказано, что задача о сбалансированной двудольной клике не аппроксимируема в полиномиальное время с

мультипликативной ошибкой r, для некоторой константы r>0 (в предположении $P \neq NP$).

2. В работе [2] доказано, что задача о сбалансированной двудольной клике не аппроксимируема в полиномиальное время с мультипликативной ошибкой $O(n^a)$, для некоторой константы $a > 0$ в предположении справедливости следующей гипотезы:

   Гипотеза. Пусть d – достаточно большая константа, не зависящая от n. Не существует полиномиального алгоритма, который отвергает почти все случайные 3-КНФ формулы с n булевыми переменными и dn скобками, причем никогда не отвергает выполнимую формулу (доля которых, как известно, стремится к нулю при достаточно большом d).

3. В работе [11] доказано, что если задача о сбалансированной двудольной клике аппроксимируема в полиномиальное время с мультипликативной ошибкой не более $2^{(\log n)^a}$, для любого $a > 0$, то задача 3-ВЫПОЛНИМОСТЬ может быть решена за время $2^{n^{3/4+\varepsilon}}$ для любого $\varepsilon > 0$.

3. Основной результат

В этой же работе доказано, что если задача о сбалансированной двудольной клике аппроксимируема в полиномиальное время с мультипликативной ошибкой не более некоторой константы, то задача о максимальной клике в графе может быть аппроксимирована в полиномиальное время с ошибкой не более $n/2^{c\sqrt{\log n}}$ для некоторой константы $c > 0$.

Нами доказана теорема о трудности аппроксимации задачи о сбалансированной двудольной клике в предположении о трудности нахождения "спрятанной большой клики" в случайном графе (см., например, [12, 13]).

Далее мы даем необходимые определения и формулируем результат.

Обозначим через $G_{n,1/2}$ случайный граф, в котором все ребра появляются независимо с вероятностью 1/2. Дадим формулировку задачи о спрятанной клике в случайном графе.

3.1 Спрятанная k-клика

Дан случайный граф $G \in G_{n,1/2}$, выбираем в нем случайное подмножество из k вершин и соединяем их ребрами, образуя полный подграф (клику). Требуется найти спрятанную клику.

Известно, что с вероятностью стремящейся к единице при $n \to \infty$ граф $G \in G_{n,1/2}$ не содержит клик размера больше $2\log n$, и максимальная клика имеет размер $(2 + o(1))\log n$. Однако, неизвестно полиномиального

алгоритма нахождения клики размера $c\log n$ при $c > 1$. В [7] даже высказано предположение о том, что задача нахождения такой клики вычислительно трудна. Косвенное подтверждение получено в [9] для одного класса популярных алгоритмов (алгоритмов, построенных на эвристиках "моделирования отжига").

В настоящее время, несмотря на довольно интенсивные исследования, неизвестно полиномиального алгоритма решения задачи о спрятанной k-клике при $k = o(\sqrt{n})$ [13], что привело к тому, что была сформулирована гипотеза о трудности ее решения (чем больше параметр k, тем сильнее эта гипотеза). Отметим, что ряд результатов уже получен некоторыми исследователями в предположении справедливости этой гипотезы [12]. Более того, она рассматривается и как один из криптографических примитивов [5].

Рассмотрим по аналогии со спрятанной кликой задачу о спрятанной двудольной клике в двудольном случайном графе. Пусть $|V_1| = |V_2| = n$. Образуем случайный двудольный (n, n)-граф следующим образом: выберем каждое ребро между V_1 и V_2 с вероятностью 1/2 независимо от других ребер. Обозначим этот класс случайных графов через $GB(n, n, 1/2)$.

3.2 Спрятанная двудольная (*k*,*k*)-клика

Дан случайный граф $G \in GB(n, n, 1/2)$, выбираем в V_1 случайное подмножество из k вершин, затем в V_2 случайное подмножество из k вершин и соединяем их ребрами, образуя полный двудольный (k, k)-подграф (двудольную клику). Требуется найти спрятанную клику.

Нетрудно показать, что с вероятностью стремящейся к единице при $n \to \infty$ граф $G \in GB(n, n, 1/2)$ не содержит двудольной (k, k)-клики с $k > 2\log n$.

Справедлива следующая

Теорема. *Пусть* $k > c\log n$, $c > 4$ *– константа. Если существует полиномиальный вероятностный алгоритм нахождения двудольной* (k, k)*-клики, спрятанной в случайном графе* $G \in GB(n, n, 1/2)$*, то существует и полиномиальный вероятностный алгоритм нахождения встроенной* $2k$*-клики в случайном графе* $G \in G_{2n,1/2}$.

Доказательство. Опишем простую сводимость задачи СПРЯТАННАЯ КЛИКА к задаче СПРЯТАННАЯ ДВУДОЛЬНАЯ КЛИКА. Итак, пусть нам дан граф $G \in G_{2n,1/2}$ содержащий клику из $2k$ вершин.

Разобъем вершины G на два подмножества V_1 и V_2, причем вершина попадает в каждый класс с вероятностью 1/2. Образуем двудольный граф $GB = (V_1, V_2, E)$, включив в E только ребра соединяющие V_1 и V_2 в G.

Довольно очевидно, что по построению полученный двудольный граф является случайным (за исключением встроенной в него $2k$-клики) с вероятностью появления ребра 1/2.

При таком разбиении вершин посмотрим, как разделились вершины $2k$-клики. Оценим снизу вероятность P_k того, что они разделились поровну, т.е. и в V_1 и в V_2 попало ровно по k вершин и, кроме того, $|V_1| = |V_2| = n$. Имеем:

$$P_k = \frac{\binom{2k}{k} \cdot \binom{2n-2k}{n-k}}{2^{2n}}.$$

Воспользуемся неравенством:

$$\binom{2m}{m} \geq c \cdot \frac{2^{2m}}{\sqrt{m}}.$$

Получим:

$$P_k \geq c^2 \frac{2^{2k} \cdot 2^{2n-2k}}{2^{2n}\sqrt{4k(n-k)}} =$$

$$= \frac{c^2}{\sqrt{4k(n-k)}} \geq \frac{c^2}{n}.$$

Отсюда сразу вытекает следствие о неаппроксимируемости задачи о максимальной сбалансированной двудольной клике в двудольном графе.

Следствие. *Пусть задача СПРЯТАННАЯ k-КЛИКА не может быть решена никаким полиномиальным вероятностным алгоритмом при* $k = \Omega(t(n))$.

Тогда для задачи о максимальной сбалансированной двудольной клике не существует полиномиального приближенного алгоритма гарантирующего мультипликативную ошибку $O(t(n))$.

В настоящее время в качестве $t(n)$ можно выбрать любую функцию $t(n) = o(\sqrt{n})$.

Список литературы

[1]. S. Khot. Improved inapproximability results for maxclique, chromatic number and approximate graph coloring, Proceedings of the 42th Annual Symposium on Foundations of Computer Science, 2001, pp. 600–609

[2]. U. Feige. Relations between average case complexity and approximation complexity, Proceedings of the 34th Annual Symposium on the Theory of Computing, 2002, pp. 534–543.

[3]. R. Peters. The maximum edge biclique problem is NP-complete, Research Memorandum 789, Faculty of Economics and Business Administration, Tilburg University, 2000.

[4]. U. Feige, R. Krauthhgamer. Finding and sertifying a large hidden clique in a semi-random graph, Random Structures and Algorithms, v. 13, 1998, pp. 457-466.

[5]. A. Juels, M. Peinado. Hiding Cliques for Cryptographic Security, Proceedings of the 9th Annual ACM-SIAM Symposium on Discrete Algorithms, 1998, pp. 678-684.

[6]. R. Karp. Reducibility among combinatorial problems, in The complexity of computer computations, Plenum Press, New York, 1972, pp. 85-103.

[7]. R. Karp. The probabilistic analysis of some combinatorial search algorithms, in Algorithms and Complexity: New directions and recent results, Academic Press, 1976, pp. 1-19.

[8]. L. Kucera. Expected complexity of graph partitioning problems, Discrete Applied Mathematics, v. 57, 1995,, pp. 193-212.

[9]. M. Jerrum. Large cliques elude the Metropolis process, Random Structures and Algorithms, v. 3, 1992, pp. 347-359.

[10]. J. Hastad. Clique is hard to approximate within $n^{1-\varepsilon}$, Proceedings of the 37th Annual IEEE Symposium on Foundations of Computing, 1997, pp. 627-636.

[11]. U. Feige, S. Kogan. Hardness of approximation of the balanced complete bipartite subgraph problem.

[12]. N. Alon, A. Andoni, T. Kaufman, K. Matulef, R. Rubinfeld, N. Xie. Testing k-wise and almost k-wise independence, Proc. Annual Symposium on the Theory of Computing, 2007, pp. 496–505.

[13]. N. Alon, M. Krivelevich, B. Sudakov. Finding a large hidden clique in a random graph, Random Structures and Algorithms, 1998, v. 13, pp. 457–466.

On the problem of finding approximation of bipatite cliques

Nikolay N. Kuzyurin
Institute for System Programming of RAS,
25, A. Solzhenitsyna str., Moscow, Russia, 109004

Abstract. In this paper, we consider the problem of finding large hidden clique in random graph and it's analog for bipartite graphs.

Keywords: random graph; large hidden clique; finding complexity

DOI: 10.15514/ISPRAS-2017-29(3)-12

For citation: Kuzyurin N.N. On the problem of finding approximation of bipatite cliques.. Trudy ISP RAN/Proc. ISP RAS, vol. 29, issue 3, 2017, pp. 225-232 (in Russian). DOI: 10.15514/ISPRAS-2017-29(3)-12

References

[1]. S. Khot. Improved inapproximability results for maxclique, chromatic number and approximate graph coloring, Proceedings of the 42th Annual Symposium on Foundations of Computer Science, 2001, pp. 600–609

[2]. U. Feige. Relations between average case complexity and approximation complexity, Proceedings of the 34th Annual Symposium on the Theory of Computing, 2002, pp. 534–543.

[3]. R. Peters. The maximum edge biclique problem is NP-complete, Research Memorandum 789, Faculty of Economics and Business Administration, Tilburg University, 2000.
[4]. U. Feige, R. Krauthhgamer. Finding and sertifying a large hidden clique in a semi-random graph, Random Structures and Algorithms, v. 13, 1998, pp. 457-466.
[5]. A. Juels, M. Peinado. Hiding Cliques for Cryptographic Security, Proceedings of the 9th Annual ACM-SIAM Symposium on Discrete Algorithms, 1998, pp. 678-684.
[6]. R. Karp. Reducibility among combinatorial problems, in The complexity of computer computations, Plenum Press, New York, 1972, pp. 85-103.
[7]. R. Karp. The probabilistic analysis of some combinatorial search algorithms, in Algorithms and Complexity: New directions and recent results, Academic Press, 1976, pp. 1-19.
[8]. L. Kucera. Expected complexity of graph partitioning problems, Discrete Applied Mathematics, v. 57, 1995,, pp. 193-212.
[9]. M. Jerrum. Large cliques elude the Metropolis process, Random Structures and Algorithms, v. 3, 1992, pp. 347-359.
[10]. J. Hastad. Clique is hard to approximate within $n^{1-\varepsilon}$, Proceedings of the 37th Annual IEEE Symposium on Foundations of Computing, 1997, pp. 627-636.
[11]. U. Feige, S. Kogan. Hardness of approximation of the balanced complete bipartite subgraph problem.
[12]. N. Alon, A. Andoni, T. Kaufman, K. Matulef, R. Rubinfeld, N. Xie. Testing k-wise and almost k-wise independence, Proc. Annual Symposium on the Theory of Computing, 2007, pp. 496–505.
[13]. N. Alon, M. Krivelevich, B. Sudakov. Finding a large hidden clique in a random graph, Random Structures and Algorithms, 1998, v. 13, pp. 457–466.