УДК 519.854

СУБЛІНІЙНИЙ ОПТИМАЛЬНИЙ НАБЛИЖЕНИЙ АЛГОРИТМ РЕОПТИМІЗАЦІЇ ДЛЯ ЗАДАЧІ ПРО МІНІМАЛЬНЕ ВЕРШИННЕ ПОКРИТТЯ ГРАФА

В.О. МИХАЙЛЮК

За наближеного розв'язання дискретних задач оптимізації виникає така ідея: чи можна, виходячи з інформації про оптимальний розв'язок екземпляра задачі (або близького до нього), використовувати цю інформацію для знаходження оптимального (або близького до нього) розв'язку екземпляра задачі, отриманого в результаті незначних локальних модифікацій вихідного екземпляра. Цей підхід, названий реоптимізацією, дозволяє в деяких випадках отримати кращу якість наближення (яке визначається як відношення значення наближеного розв'язку до точного і називається відношенням апроксимації) у локально модифікованих екземплярах, ніж у вихідних. Якщо для деяких оптимізаційних задач відношення апроксимації не можна покращити (наприклад, у класі всіх наближених алгоритмів із поліноміальною складністю), то таке відношення називають пороговим або оптимальним (алгоритм на якому досягається це відношення також називають пороговим або оптимальним). Складність алгоритмів оцінюється кількістю звернень (запитів) до спеціального оракулу. Для реоптимізації задачі про мінімальне вершинне покриття графа (за добавлення однієї вершини і деякої множини ребер) отриманий (3/2)-наближений алгоритм із адитивною помилкою з сублінійною (константною) складністю. Показано, що відношення апроксимації 3/2 є пороговим у класі алгоритмів із константною складністю.

ВСТУП

Поняття сублінійних алгоритмів виникло завдяки необхідності обробки величезних масивів інформації, що має місце в багатьох прикладненнях. Кажуть, що алгоритм має сублінійну складність, якщо час його роботи оцінюється величиною $o(n)$, де n — розмірність входу. Для того, щоб сублінійний алгоритм був точним треба, щоб він використовував паралельну (некласичну) обробку даних або враховував спеціальні обмеження на вхідні дані (наприклад, логарифмічний бінарний пошук у відсортованих масивах). Інакше такий алгоритм не в змозі прочитати весь вхід, щоб забезпечити деякий результат роботи. Тому сублінійні алгоритми мають видавати результат без читання всього входу та мати доступ до нього тільки деякими обмеженими порціями і бути випадковими, видаючи тільки наближені розв'язки. Ефективність таких алгоритмів вимірюється так званою складністю запитів (query complexity) — кількістю доступів до деякого оракулу, що обробляє вхід (порціями).

Сублінійні алгоритми знайшли широке застосування в розділі теоретичної інформатики, який називається «перевірка властивостей проблем» (property testing) [1, 2]. Перевірка властивостей проблем (релаксація проблем розпізнавання) пов'язана з розробкою алгоритмів (сублінійних), що розрізняють об'єкти із заданою властивістю від об'єктів, які далекі від цієї властивості (ε-далекі, ε-far). Наприклад, для задач теорії графів із n вершинами і обмеженою ступінню вершин d (ступінь вершини — число ребер їй інцидентних) властиво: бути ε-далекими від деякої властивості P означає добавити або видалити εdn ребер, щоб граф мав цю властивість P. Сублінійні алгоритми перевірки властивостей (тестери) використовують у теорії навчання і теорії наближення.

Задачу про мінімальне вершинне покриття графа (Min-Vertex-Cover) сформулюємо так. Вершинним покриттям (vertex cover) неорієнтованого графа $G=(V,E)$ називається така підмножина його вершин $V'\subset V$, що для довільного ребра $(u,v)\in E$ хоча б одна з вершин u або v міститься в V'. Під розміром покриття розуміється число його елементів. Вимагається знайти покриття мінімального розміру. Через VC_G будемо позначати оптимальний розв'язок цієї задачі, а через C' — наближений, $|VC_G|$ — число елементів у множині VC_G (розмір), n число вершин в графі.

Задача про мінімальне вершинне покриття графа з обмеженою ступінню $(d=\text{const})$ у контексті перевірки властивостей визначає чи існує сублінійний алгоритм, що з високою ймовірністю встановлює чи має граф вершинне покриття розміру $\rho n,$ або цей граф є ε-далеким від цієї властивості. Останнє означає, що не менше ніж $\varepsilon\cdot dn$ ребер треба виключити з графа, щоб він мав вершинне покриття розміром $\rho n.$ У [2] відмічено, що результати по NP-складності наближення мінімального вершинного покриття розповсюджуються і на цю задачу.

У випадку графів із обмеженою ступінню постановка задачі сублінійного наближення мінімального вершинного покриття є наступне розширення задачі перевірки властивостей. Мета цієї задачі полягає у розрізненні випадку, коли граф має вершинне покриття розміру $\rho n,$ і випадку, коли граф є ε-далеким від властивості мати вершинне покриття розміром $\alpha\cdot\rho n.$ У [3] показано, що будь-який наближений алгоритм для задачі про мінімальне вершинне покриття з мультиплікативною помилкою не більше 7/6 вимагає $\Omega(n)$ запитів.

При наближеному розв'язанні дискретних задач оптимізації виникає така ідея: виходячи з оптимального розв'язку екземпляра задачі (або близького до нього) чи можна використати цю інформацію для знаходження оптимального (або близького до нього) розв'язку екземпляра проблеми, отриманого в результаті незначних локальних модифікацій вихідного екземпляра. Цей підхід названо реоптимізацією (вперше запропоновано в [4]) дозволяє, наприклад, у деяких випадках отримати кращу якість наближення (яке визначається як відношення значення наближеного розв'язку до точно модифікованих екземплярах, ніж у вихідних. Якщо для деяких оптимізаційних задач відношення апроксимації не можна покращити (наприклад, у класі всіх наближених алгоритмів із поліноміальною складністю), то таке відношення називають пороговим або оптимальним (алгоритм на якому досягається це відношення також називають пороговим або оптимальним). У [5] вдалося отримати точне значення порогового відношення апроксимації для реоптимізації узагальнених проблем про виконуваність спеціального класу для алгоритмів поліноміальної складності.

Мета роботи — отримання точного значення порогового відношення апроксимації для реоптимізації задачі про мінімальне вершинне покриття в графі в класі сублінійних алгоритмів, зокрема константної складності.

ОСНОВНІ ОЗНАЧЕННЯ І ПОНЯТТЯ

Розглянемо задачу про мінімальне вершинне покриття графа $G=(V,E)$. Діаметр графа — це максимальна з відстаней між парами його вершин. Відстань між вершинами визначається як найменше число ребер, які необхідно пройти, щоб дістатися з однієї вершини в іншу. Інакше кажучи, діаметр — це відстань між двома вершинами графа, максимально віддаленими одна від одної.

Означення 1. Нехай $\alpha \geq 1$ та $0 \leq \varepsilon \leq 1$. Значення C' є (α,ε) — наближенням $|VC_G|$, якщо $|VC_G| \leq C' \leq \alpha |VC_G| + \varepsilon n$. Алгоритм, який для даного $\varepsilon > 0$ в якості вхідного параметра обчислює з ймовірністю не меншою ніж 2/3 (α,ε)-наближення $|VC_G|$ для деякого значення α, називається (α,ε)-наближеним алгоритмом або α-наближеним алгоритмом із адитивною помилкою.

Складність алгоритмів будемо оцінювати кількістю звернень (запитів) до спеціального оракулу. Два ребра є сусідніми, якщо вони мають спільну вершину (дві вершини є сусідніми, якщо вони з’єднані ребром). На запит (v,i) відповіддю є i-й сусід вершини v або спеціальний символ у випадку, коли v має менше ніж i сусідів. Іноді дозволяються оракули визначення ступені вершин, коли для довільної вершини v видається її ступінь.

Означення 2. Будемо говорити, що для задачі про мінімальне вершинне покриття графа існує оптимальний (α,ε)-наближений алгоритм із константною складністю, якщо виконуються такі умови:

- для довільного $\varepsilon > 0$ для задачі існує (α,ε)-наближений алгоритм із константною складністю;
- для довільного $\varepsilon > 0$ існує таке $\delta > 0$, що для задачі не існує $(\alpha-\varepsilon,\delta)$-наближеного алгоритму з константною складністю.

При цьому α називається *порогом відношення* апроксимації наближеного алгоритму, а сам алгоритм — оптимальним α-наближеним алгоритмом із адитивною помилкою з константною складністю.

Будемо використовувати поняття розподіленого алгоритму (distributed algorithm) [6, 7]. В основі розподіленої обчислювальної моделі лежить синхронна мережа G, в якій вершини представляють процесори, а ребра — комунікаційні канали. Один раунд розподіленого алгоритму полягає в пересилці повідомлень від кожної вершини до сусідніх вершин. Після k раундів кожна вершина закінчує свої обчислення і вся мережа досягає деякої глобальної мети. Наприклад, якщо мета полягає в обчисленні вершинного покриття, кожна вершина має вирішити чи належить вона до вершинного покриття. У локальній обчислювальній моделі кожна вершина виконує своє завдання, виходячи з локальної інформації, якщо k менше діаметра графа.

СУБЛІНІЙНЕ (КОНСТАНТНЕ) НАБЛИЖЕННЯ РОЗМІРУ МІНІМАЛЬНОГО ВЕРШИННОГО ПОКРИТТЯ ГРАФУ

У [6] показано зведеність від локальних розподілених наближених алгоритмів (distributed algorithms) до сублінійних наближених алгоритмів (константної складності). Коротко опишемо суть результатів.

Наведемо розподілений алгоритм для мінімального вершинного покриття графа. Для ребра e позначимо $\Gamma(e)$ множину його сусідів та $d(e)$ розмір $\Gamma(e)$. Наведемо випадковий розподілений алгоритм, що знаходить вершинне покриття розміром не більше ніж у $(2+\delta)$ раз більше від мінімального з високою константною ймовірністю.

Алгоритм 1 (розподілене наближення мінімального вершинного покриття)**.**

1. Кожне ребро активізує само себе.
2. Для i від 1 до $r = \Theta(\log(d/\delta))$ виконати:

а) кожне активоване ребро e вибирає себе з ймовірністю $\frac{1}{4d(e)}$. Якщо $d(e)=0$, то e вибирається з ймовірністю 1;

б) кожні два сусідніх ребра, що були вибрані, відміняють свій вибір;

в) кожна вершина інцидентна вибраному ребру додає себе до вершинного покриття;

г) вибрані ребра і сусіди вибраних ребер дезактивують себе;

д) активні ребра оновлюють свої степені, щоб вони були рівні числу своїх активних сусідів.

3. Одна вершина кожного раніше активованого ребра додає себе до вершинного покриття.

Має місце наступне твердження.

Теорема 1 [6]**.** Для довільного $\delta > 0$ та кожного графа $G=(V,E)$ з максимальною степінню d алгоритм 1 знаходить вершинне покриття $C \subseteq V$ таке, що з ймовірністю не меншою ніж 5/6 $|VC_G| \le |C| \le$ $\le (2+\delta)|VC_G|$. При цьому складність алгоритму є $d^{O(\log(d/\delta))}$.

Алгоритм 2 (сублінійне наближення для VC_G)**.**

1. Рівномірно і незалежно генерується $s = 4/\delta^2$ вершин графа G. Нехай S множина цих вершин.

2. Для кожного $v \in S$ розглядається підграф $G_r(v)$ індукований $(r+1)$-сусідством v, де $r = \Theta\left(\log\left(\frac{d}{\delta}\right)\right)$, як в алгоритмі 1.

3. Для $\bigcup_{v \in S} G_r(v)$ застосовується алгоритм 1 (у послідовному порядку). Для кожного $v \in S$ нехай $\chi_v = 1$, якщо алгоритм додає v до покриття, інакше — $\chi_v = 0$.

4. Виведення $\overline{VC} = \frac{n}{s} \sum_{v \in S} \chi_v$.

Встановлено такий результат.

Теорема 2 [6]**.** Для довільного $\delta > 0$ та кожного графу G алгоритм 2 виводить із ймовірністю не меншою ніж 2/3 оцінку $\overline{VC}$, що задовольняє

$$|VC_G| - \delta n \le \overline{VC} \le 2 \cdot |VC_G| + \delta n .$$

Складність алгоритму є $d^{O(\log(d/\delta))}$.

Таким чином, алгоритм 2 є $(2, \delta)$-наближеним алгоритмом для Min-Vertex-Cover.

У [7] встановлено таку нижню оцінку якості наближення.

Теорема 3 [7]**.** Для довільних двох даних констант γ та ε знайдеться константа d така, що визначення з константною ймовірністю $(2-\gamma, \varepsilon)$-наближення мінімального вершинного покриття графа з n вершинами ступені d вимагає $\Omega(n)$ запитів.

З теореми 3 випливає, що не існує β-наближеного сублінійного алгоритму (з константною складністю) для задачі Min-Vertex-Cover із відношенням апроксимації β строго меншим 2. З врахуванням означення 2 отримаємо твердження.

Наслідок 1. Для задачі Min-Vertex-Cover існує оптимальний 2-наближений алгоритм із адитивною помилкою з константною складністю.

ПОРІГ ВІДНОШЕННЯ АПРОКСИМАЦІЇ НАБЛИЖЕНОГО АЛГОРИТМУ З КОНСТАНТНОЮ СКЛАДНІСТЮ ДЛЯ РЕОПТИМІЗАЦІЇ MIN-VERTEX-COVER

Приведемо реоптимізаційний варіант для задачі Min-Vertex-Cover. Нехай $I = G = (V, E)$ — довільний екземпляр задачі Min-Vertex-Cover, екземпляр $I' = G' = (V', E')$ отримується з I добавленням довільної вершини v_{new} ($V' = V \cup \{v_{\text{new}}\}$), з деякою множиною ребер $\{e_{\text{new}}\}$ інцидентних v_{new} ($E' = E \cup \{e_{\text{new}}\}$). Через $N(v)$ будемо позначати множину всіх сусідніх вершин вершини $v \in V$.

Задача Ins-Min-Vertex-Cover. Вхідні дані. Екземпляр I задачі Min-Vertex-Cover, оптимальний розв'язок VC_G екземпляра I.

Результат. Використовуючи VC_G, знайти оптимальний розв'язок $VC_{G'}^*$ екземпляра I', описаного вище.

Мета. Мінімізувати число вершин у вершинному покритті $I' = G'$.

Теорема 4. Якщо для задачі Min-Vertex-Cover існує (ρ,δ)-наближений алгоритм із константною складністю, то для задачі Ins-Min-Vertex-Cover (реоптимізація Min-Vertex-Cover) існує $(\varphi(\rho),\delta_1)$-наближений алгоритм із константною складністю, де $\varphi(\rho) = 2 - 1/\rho$, а $\delta_1 = \delta/\rho$.

Доведення. Якщо $VC_{G'}^*$ містить v_{new}, то $VC_G \cup \{v_{\text{new}}\}$ оптимальний розв'язок I'.

Нехай $VC_{G'}^*$ не містить v_{new}. Щоб бути допустимим, $VC_{G'}^*$ має містити всі сусідні вершини $N(v_{\text{new}})$ і, отже, $VC_G \cup N(v_{\text{new}})$ — допустимий розв'язок I'. Оскільки $|VC_G| \le |VC_{G'}^*|$, то

$$\left|VC_G \cup N(v_{\text{new}})\right| \le |VC_{G'}^*| + \left|N(v_{\text{new}})\right|. \qquad (1)$$

Побудуємо ще один допустимий розв'язок VC^1:

а) видалимо v_{new} і всі сусідні вершини $N(v_{\text{new}})$;

б) застосуємо (ρ,δ)-наближений алгоритм із константною складністю до вершин, що залишилися;

в) добавимо вершини $N(v_{\text{new}})$.

У результаті отримаємо:

$$|VC^1| \le \rho(|VC_{G'}^*| - |N(v_{\text{new}})|) + \delta n + |N(v_{\text{new}})| =$$

$$= \rho|VC_{G'}^*| - (\rho-1)|N(v_{\text{new}})| + \delta n. \qquad (2)$$

Помножимо (1) на $\rho - 1$ та складаючи з (2), отримаємо:

$$(\rho-1)\left|VC_G \cup N(v_{\text{new}})\right| + |VC^1| \le (\rho-1)|VC_{G'}^*| + (\rho-1)\left|N(v_{\text{new}})\right| +$$

$$+ \rho|VC_{G'}^*| - (\rho-1)\left|N(v_{\text{new}})\right| + \delta n = (2\rho-1)|VC_{G'}^*| + \delta n.$$

Звідси $(\rho - 1 + 1)\min\{\left|VC_G \cup N(v_{new})\right|, |VC^1|\} \le (2\rho-1)|VC_{G'}^*| + \delta n$.

З розв'язків $VC_G \cup N(v_{\text{new}})$ та VC^1 виберемо кращий, тобто з меншим числом елементів і позначимо його $\overline{VC}$. Остаточно отримаємо $|\overline{VC}| \le \dfrac{2\rho-1}{\rho}|VC_{G'}^*| + \dfrac{\delta}{\rho}n$ або $|\overline{VC}| \le \varphi(\rho)|VC_{G'}^*| + \delta_1 n$, де $\varphi(\rho) = 2 - 1/\rho$, а $\delta_1 = \delta/\rho$. Таким чином, для для задачі Ins-Min-Vertex-Cover отримали $(\varphi(\rho),\delta_1)$-наближений алгоритм із константною складністю.

Наслідок 2. Для задачі Ins-Min-Vertex-Cover існує (3/2)-наближений алгоритм із адитивною помилкою з константною складністю.

Доведення. Згідно з теоремою 2 (або із наслідком 1) треба покласти в умовах теореми 4 $\rho = 2$ і застосувати в доведенні алгоритм 2 в п. б).

Теорема 5. Нехай A — будь-який γ-наближений алгоритм із адитивною помилкою з константною складністю для задачі Ins-Min-Vertex-Cover, тоді $\gamma \geq 3/2$.

Доведення. Згідно з теоремою 4, маючи (ρ, δ)-наближений алгоритм із константною складністю для задачі Min-Vertex-Cover, отримаємо $(\varphi(\rho), \delta_1)$-наближений алгоритм із константною складністю для задачі Ins-Min-Vertex-Cover, де $\varphi(\rho) = 2 - 1/\rho$, а $\delta_1 = \delta/\rho$. Отже маємо $\varphi(\rho)$-наближений алгоритм із адитивною помилкою з константною складністю для задачі Ins-Min-Vertex-Cover. Доведення теореми проведемо від супротивного. Нехай $\gamma = \varphi(\rho) < 3/2 = \varphi(2)$. Оскільки функція $\varphi(\rho)$ є зростаючою функцією свого аргументу, звідси будемо мати $\rho < 2$. Таким чином, у теоремі 4 застосовувався ρ-наближений алгоритм із константною складністю при $\rho < 2$, що суперечить теоремі 3 (та наслідку 1).

Теорема 6. Для задачі Ins-Min-Vertex-Cover існує оптимальний (3/2)-наближений алгоритм із адитивною помилкою з константною складністю.

Доведення слідує з теорем 4 та 5.

Зауваження. Відмітимо, що якість наближення алгоритму з теореми 6 краща (менша) ніж якість наближення (2) оригінального алгоритма задачі Min-Vertex-Cover.

ВИСНОВКИ

Як правило знаходження порогових значень відношень апроксимації (зокрема нижніх оцінок) наближених алгоритмів розв'язування дискретних задач оптимізації здійснюється за допомогою деяких загально прийнятих гіпотез теоретичної інформатики. Наприклад, завдяки гіпотезі співвідношення класів складності задач по включенню $(P \neq NP)$ вдалося отримати порогові значення апроксимації для наближених алгоритмів широкого класу задач про узагальнену виконуваність [8]. Завдяки унікальній ігровій гіпотезі (Unique Games Conjecture, UGC) вдалося розповсюдити ці результати на інші класи задач (наприклад, Max Cut) [9,10]. Подібні результати (отримання нижніх оцінок відношення апроксимації завдяки використанню гіпотез) називають результатами по умовній неапроксимованості. У цій роботі наведено приклад результатів з безумовної неапроксимованості для реоптимізації задачі про мінімальне вершинне покриття графа.

ЛІТЕРАТУРА

1. *Goldreich O., Goldwasser S., Ron D.* Property testing and its connection to learning and approximation // Journal of the ACM. — 1998. — **45**(4). — P. 653–750.
2. *Goldreich O., Ron D.* Property testing in bounded degree graphs // Algorithmica. — 1997. — **32**(2). — P. 302–343.

3. *Bogdanov A., Obata Kenji, Trevisan L.* A lower bound for testing 3-colorability in bounded-degree graphs // In Proceedings of the Forty-Third Annual Symposium on Foundations of Computer Science. — 2002. — P. 93–102.
4. *Archetti C., Bertazzi L., Speranza M.G.* Reoptimizing the traveling salesman problem // Networks. — 2003. — **42** (3). — P. 154–159.
5. *Михайлюк В.А., Сергиенко И.В.* Реоптимизация обобщенных проблем о выполнимости с аппроксимационно-устойчивыми предикатами // Кибернетика и системый анализ. — 2012. — **48**, № 1. — С. 89–104.
6. *Marko S., Ron D.* Distance approximation in bounded-degree and general sparse graphs // In Proceedings of the Tenth International Workshop on Randomization and Computation (APPROX-RANDOM). Lecture Notes in Computer Science. — 2006. — 4110. — P. 475–486
7. *Parnas M., Ron D.* Approximating the minimum vertex cover in sublinear time and a connection to distributed algorithms // Theoretical Computer Science. — 2007. — 381(1–3). — P.183–196.
8. *Hastad J.* Some optimal inapproximability results // Journal of the ACM. — 2001. — **48**, № 4. — P. 798–859.
9. *Khot S., KIndler G., Mossel E., O'Donnell R.* Optimal Inapproximability Results for Max-Cut and Other 2-Variable CSPs? // In FOCS. — 2004. — P.146–154.
10. *Khot S.* On the Unique Games Conjecture // In Proc. of the 25-th Annual IEEE Conference on Computational Complexity. — 2010. — P. 99–121.

Поступила 25.05.2012