Сравнительный анализ методов оценивания бикликового покрытия

Попов Максим
6 Курс ФИВТ
Кафедра: Анализ данных

23 июня 2016

Определения и понятия

Определение

Бикликой *неориентированного графа называется подмножество его вершин, образующих полный двудольный подграф.*

Определение

Бикликовым покрытием $\boldsymbol{bcc(G)}$ *графа* G *будем называть наименьшее число, возможно, пересекающихся биклик, которыми можно покрыть все ребра графа* G.

Мотивировка

1) Естественный вопрос теории графов + приложения в вычислительных задачах.
2) Приложения в коммуникационной сложности:
   - $bcc(G)$ дает **нижнюю** оценку для **детерминированной** коммуникационной сложности;
   - $bcc(G)$ дает **точную** оценку для **недетерминированной** коммуникационной сложности.

Методы оценивания

Нас интересует нижняя оценка величины $bcc(G)$.

Сравниваем 3 метода доказательства оценок:

1) **Трудного множества**;
2) **Куликова-Юкны**;
3) **Информационных неравенств**.

Основные результаты

1) Сравнил методы оценивания на графах, построенных при помощи геометрических конфигураций;
2) Доказал, что метод трудного множества для произвольного графа дает оценки не хуже метода Куликова-Юкны;
3) С помощью метода случайных графов показал, что техника трудного множества может давать оценки значительно сильнее, чем метод Куликова-Юкны;
4) Получил обобщение метода информационных неравенств для задачи коммуникационной сложности с $m > 2$ участниками.

Метод трудного множества

Определение

В графе G подмножество рёбер S называем **трудным**, *если для любых двух рёбер $(x, y) \in S$ и $(x', y') \in S$*

- *либо (x, y') не является ребром,*
- *либо (x', y) не является ребром.*

$\boldsymbol{fool(G)}$ – *размер максимального трудного множества.*

Теорема

Если S трудное множество графа G, то $bcc(G) \geq |S|$. В частности,

$$bcc(G) \geq fool(G).$$

Метод Куликова-Юкны

Jukna S., Kulikov A. S. On covering graphs by complete bipartite subgraphs. 2009.

Теорема

Для произвольного графа $G = (V, E)$

$$bcc(G) \geq \frac{\nu(G)^2}{|E|},$$

где $\nu(G)$ – размер максимального паросочетания в G.

Метод информационных неравенств

Kaced T., Romashchenko A.E., Vereshchagin N.K. Conditional Information Inequalities and Combinatorial Applications. 2015.

Теорема

Пусть ребра двудольного графа $G = (L, R, E)$ раскрашены по следующему правилу:

(*) *для произвольной биклики $C \subseteq G$ и для произвольной пары ребер (x, y') и (x', y) из C одного цвета a, цвет ребра (x, y) тоже a.*

Зададим распределение вероятностей на ребрах:

- $X =$ [*левый конец ребра*]; $Y =$ [*правый конец ребра*];
- $A =$ [*цвет ребра*].

Тогда $bcc(G) \geq 2^{\frac{1}{2}(H(A|X)+H(A|Y)-H(A))}$.

Трудное множество vs оценка Куликова-Юкны

Теорема (FS всегда не слабее KJ)

В любом графе $G = (V, E)$ среди ребер максимального паросочетания найдется трудное множество размера $\geq \frac{\nu(G)^2}{|E|}$.

Теорема (FS иногда значительно сильнее KJ)

Для произвольных $\alpha \in [0, \frac{1}{3})$ и $\beta \in (\alpha, \frac{1+\alpha}{2})$ при достаточно больших n существует двудольный граф $G = (L, R, E)$ такой, что $|L| = |R| = n$, на котором метод трудного множества дает оценку хотя бы n^β, а оценка Куликова-Юкны не превосходит $n^\alpha + o(1)$.

Обобщение информационного метода

Теорема

Пусть ребра гиперграфа $G = (X_1, X_2, \ldots, X_m, E)$ *раскрашены по следующему правилу:*

(*) *для произвольного полного m-дольного гиперграфа* $C \subseteq G$ *и для произвольного набора ребер* $(x_{1,1}, \ldots, x_{1,m}), \ldots, (x_{m,1}, \ldots, x_{m,m})$ *одного цвета a, цвет ребра* $(x_{1,1}, x_{2,2}, \ldots, x_{m,m})$ *тоже a.*

Зададим распределение вероятностей на ребрах:

- $X_i = [i$*-ая вершина ребра*$]$;
- $A = [$*цвет ребра*$]$.

Тогда выполняется неравенство:

$$bcc(G) \geq 2^{\frac{1}{m}(H(A|X_1)+\ldots+H(A|X_m)-(m-1)H(A))}$$

Используемые теоремы и техники

1) При доказательстве первой теоремы была придумана конструкция графа четырехсторонников $\widetilde{G}$. Доказано, что $fool(G) = \omega(\widetilde{G})$, $\max\limits_{K_{r,s} \subseteq G} \{r \cdot s\} = \alpha(\widetilde{G})$ и $bcc(G) = \chi(\widetilde{G})$. Окончательное утверждение теоремы получается применением теоремы Турана к графу $\widetilde{G}$.
2) Во второй теореме использовались случайные графы Эрдеша-Реньи. При оценивании максимального паросочетания применялась лемма Холла, а при оценивании количества ребер – неравенство Хефдинга.
3) В последней теореме использовалось не Шенноновское информационное неравенство:

$$H(A|X_1, B) + H(A|X_2, B) + \ldots + H(A|X_m, B) \leq (m-1)H(A|B)$$

Открытые вопросы

1) Не удалось доказать, что метод трудных множеств работает **почти наверное** (для случайных графов) намного лучше, чем оценка Куликова-Юкны. А именно не получилось доказать

$$\sum_{T:T\sim S_0} P\{I_k(T)=1 \mid I_k(S_0)=1\} = o(\mathbb{E}[f_k(G)])$$

2) Не получилось сравнить в общем случае метод трудных множеств с методом информационных неравенств.

Спасибо за внимание!