

УДК 519.1

О. И. ДУГИНОВ¹, И. Г. КУЗНЕЦОВА²

ЗАДАЧА МИНИМАЛЬНОГО ПОПОЛНЕНИЯ ДВУДОЛЬНОГО ГРАФА

(Представлено академиком И. В. Гайшуном)

¹Институт математики НАН Беларуси, Минск, Беларусь
oduginov@gmail.com²Белорусский государственный университет, Минск, Беларусь
irene.kuzn@gmail.com

Рассматривается графовая задача, в которой задан двудольный граф с выделенной долей и требуется добавить в граф наименьшее число дополнительных ребер так, что множество вершин выделенной доли получившегося графа можно разбить на заданное число непустых множеств, каждое из которых содержит только вершины с одинаковыми окружениями. В работе установлено, что задача является NP -трудной в классе P_4 -свободных двудольных графов и предлагается алгоритм, который решает задачу в классе $2K_2$ -свободных двудольных графов.

Ключевые слова: пополнение двудольного графа, классы графов, вычислительная сложность.

O. I. DUGINOV¹, I. G. KOUZNETSOVA²

CLUSTERING MINIMUM BICLIQUE COMPLETION OF A BIPARTITE GRAPH

¹Institute of Mathematics of the National Academy of Sciences of Belarus, Minsk, Belarus
oduginov@gmail.com²Belarusian State University, Minsk, Belarus
irene.kuzn@gmail.com

In this article we show that the clustering minimum biclique completion problem is NP -complete in the class of P_4 -free bipartite graphs. We have also proposed a dynamic programming algorithm for that problem restricted to $2K_2$ -free bipartite graphs.

Keywords: clustering minimum biclique completion problem, graph classes, computational complexity.

В настоящей работе изучается следующая задача: для заданных двудольного графа с выделенной долей и натурального числа p требуется найти минимальное число добавлений в граф отсутствующих в нем ребер (между вершинами из разных долей) так, что множество вершин выделенной доли этого графа можно разбить ровно на p непустых множеств с условием, что любые две вершины выделенной доли, попадающие в одно множество разбиения, имеют одинаковые окружения. Эта задача принадлежит классу задач модификации графа и возникает в области телекоммуникационных сетевых технологий.

Теоретико-графовая терминология и предварительные сведения. Рассматриваются только неориентированные графы $G = (V, E)$ без кратных ребер и петель с множеством вершин $V = VG$ и множеством ребер $E = EG$. Используется стандартная теоретико-графовая терминология (см., напр., [1]). Множество вершин графа G , смежных с вершиной $v \in V$, называется *окружением* вершины v и обозначается через $N(v)$. Пусть U – подмножество вершин графа G . Тогда множество всех вершин графа, смежных хотя бы с одной вершиной из U , называется *окружением* множества U и обозначается через $N(U)$. Подграф графа G , порожденный множеством вершин U , обозначается через $G[U]$. Пусть H – некоторый граф. Граф G называется H -свободным, если он не содержит порожденных подграфов, изоморфных графу H .

Граф $G = (V, E)$ называется *двудольным*, если его множество вершин может быть разбито на два подмножества $V = X \cup Y$ таким образом, что каждое ребро графа G соединяет вершины из разных подмножеств X, Y . Множества X, Y называют *долями* графа G ; используют обозначение $G = (X \cup Y, E)$. Двудольный граф $G = (X \cup Y, E)$ называется *полным двудольным*, если каждая вершина из множества X соединена ребром с каждой вершиной из множества Y , т. е. $E = \{\{x, y\} : x \in X, y \in Y\}$.

Любой двудольный граф $G = (X \cup Y, E)$ может быть дополнен до полного двудольного графа путем добавления в него недостающих ребер, т. е. ребер из множества $\{\{x, y\} : x \in X, y \in Y\} \setminus E$. Число ребер, которое необходимо при этом добавить к графу G , будем называть *недостатком* графа G и обозначать как $c(G)$. Таким образом, $c(G) = |X| \cdot |Y| - |E|$.

Пусть $G = (X \cup Y, E)$ – двудольный граф с непустой долей X и p – натуральное число, не превосходящее $|X|$. Рассмотрим разбиение множества X

$$X = X_1 \cup X_2 \cup \dots \cup X_p,$$

где $X_i \cap X_j = \emptyset$ при $i \neq j$, на p непустых множеств X_1, X_2, \dots, X_p . *Стоимостью* разбиения будем называть сумму недостатков двудольных графов $G[X_i \cup N(X_i)]$, $i = \overline{1, p}$. Так как граф G двудольный с долями X, Y и семейство множеств X_1, X_2, \dots, X_p образует разбиение множества X , то

$$\sum_{i=1}^p c(G[X_i \cup N(X_i)]) = \sum_{i=1}^p |X_i| \cdot |N(X_i)| - |E|.$$

Величиной разбиения множества X на p непустых множеств X_1, X_2, \dots, X_p будем называть сумму $\sum_{i=1}^p |X_i| \cdot |N(X_i)|$. Таким образом, стоимость разбиения множества X на p множеств меньше величины этого разбиения ровно на число ребер в графе G . Минимальную стоимость (величину) разбиения множества X на p множеств будем обозначать через $\xi_p(G)$ (соответственно, $\zeta_p(G)$). Легко видеть, что

$$\xi_p(G) = \zeta_p(G) - |E|. \quad (1)$$

С параметром $\xi_p(G)$ связана задача, которая в виде задачи распознавания свойств формулируется следующим образом:

Пополнение до разбиения на биклики

Условие: Заданы двудольный граф $G = (X \cup Y, E)$, натуральные числа p, k .

Вопрос: Существует ли разбиение множества X на p непустых множеств, стоимость которого не превосходит k ?

Отметим, что сформулированный выше вопрос равносильен следующему: верно ли, что $\xi_p(G) \leq k$? В оптимизационной версии этой задачи для заданных двудольного графа $G = (X \cup Y, E)$ и натурального числа p требуется найти разбиение наименьшей стоимости множества X на p непустых множеств. Задача NP -полна даже для фиксированного $p = 2$ [2]. Эта задача, по-видимому, впервые была сформулирована в диссертации Н. Форе [3] и широко изучается в ряде работ [2–6]. Отметим также, что взвешенная версия этой задачи изучается в [7].

Цель настоящей работы – изучение вычислительной сложности рассматриваемой задачи в подклассах двудольных графов. Мы покажем, что задача является NP -полной в классе P_4 -свободных двудольных графов (здесь P_4 – простая цепь на четырех вершинах) и дадим алгоритм с трудоемкостью $O(p|X|^4|Y|)$, который находит разбиение наименьшей стоимости множества X на p непустых множеств для $2K_2$ -свободного двудольного графа $G = (X \cup Y, E)$ (здесь граф $2K_2$ – граф с множеством вершин $\{a, b, c, d\}$ и множеством ребер $\{\{a, b\}, \{c, d\}\}$).

В виду тесной связи параметров $\xi_p(G)$, $\zeta_p(G)$ целесообразно рассмотреть задачу, связанную с параметром $\zeta_p(G)$, которая в виде задачи распознавания формулируется следующим образом:

Мин. сумма произведений мощностей множеств и их окружений

Условие: Заданы двудольный граф $G = (X \cup Y, E)$ и натуральные числа p, k .

Вопрос: Существует ли разбиение множества X на p непустых подмножеств, величина которого не превосходит k ? Эквивалентно, верно ли, что $\zeta_p(G) \leq k$?

Нетрудно видеть, что эта задача принадлежит классу NP . Из равенства (1) следует, что задача «Мин. сумма произведений мощностей множеств и их окружений» и задача «Пополнение до разбиения на биклики» полиномиально изоморфны, что означает, по определению, существование такой биективной функции f , что f реализует полиномиальное сведение первой задачи ко второй, а обратная функция f^{-1} реализует полиномиальное сведение второй задачи к первой (см. [8, с. 200]).

NP-полнота задач, ограниченных классом P_4 -свободных двудольных графов. В этом разделе покажем, что задача «Мин. сумма произведений мощностей множеств и их окружений» является NP-полной даже будучи ограниченной классом P_4 -свободных двудольных графов.

Напомним, что P_4 -свободные двудольные графы – двудольные графы, не содержащие порожденных подграфов, изоморфных простой цепи P_4 . Известно, что структура P_4 -свободных двудольных графов проста, а именно [9, следствие 2.2], двудольный граф является P_4 -свободным тогда и только тогда, когда каждая его компонента связности изоморфна полному двудольному графу (возможно с пустой одной долей).

Т е о р е м а 1. *Задача «Мин. сумма произведений мощностей множеств и их окружений» в классе P_4 -свободных двудольных графов является NP-полной.*

Д о к а з а т е л ь с т в о. Рассмотрим задачу «Мин. сумма произведений мощностей множеств и их окружений» в случае, когда заданный двудольный граф $G = (X \cup Y, E)$ является P_4 -свободным. Заметим, что если число p не меньше чем число компонент связности в графе G , то минимальная величина разбиения множества X на p непустых множеств равна $|E|$. Мы ограничимся только случаем, когда число p меньше чем число компонент связности графа G .

Нетрудно видеть, что задача «Мин. сумма произведений мощностей множеств и их окружений», ограниченная классом P_4 -свободных двудольных графов, принадлежит классу NP. Построим полиномиальное сведение от специального NP-полного случая задачи

4-Разбиение

Условие: Заданы множество $A = \{a_1, a_2, \dots, a_{4p}\}$, элементами которого являются натуральные числа, и натуральное число b , причем

$$\frac{b}{5} < a_i < \frac{b}{3}, \forall i \in \{1, 2, \dots, 4p\}$$

и

$$\sum_{i=1}^{4p} a_i = pb. \quad (2)$$

Вопрос: Можно ли множество индексов $I = \{1, 2, \dots, 4p\}$ разбить на p непустых множеств I_1, I_2, \dots, I_p так, что

$$\sum_{i \in I_j} a_i = b \quad (3)$$

для каждого $j \in \{1, 2, \dots, p\}$?

Известно [8, с. 125], что эта задача NP-полна даже в случае, когда значения всех числовых параметров ограничены сверху числом $2^{16} |A|^4 = 2^{24} p^4$.

Построим полиномиальное сведение от задачи «4-разбиение» с ограниченными числовыми параметрами. По множеству A построим двудольный граф $G = (X \cup Y, E)$, который состоит из $4p$ компонент связности G_1, G_2, \dots, G_{4p} , и при этом каждая компонента связности G_i изоморфна полному двудольному графу, обе доли которого содержат по a_i вершин. Положим $k = pb^2$. Отметим, что граф G является P_4 -свободным двудольным графом и может быть построен за полиномиальное время в виду ограниченности элементов множества A .

Утверждается, что существует разбиение множества индексов $I = \{1, 2, \dots, 4p\}$

$$I = I_1 \cup I_2 \cup \dots \cup I_p \quad (4)$$

такое, что для каждого множества $I_j, j \in \{1, 2, \dots, p\}$, выполняется равенство (3) тогда и только тогда, когда $\zeta_p(G) \leq k$.

Пусть существует такое разбиение (4). Рассмотрим разбиение доли X графа G на p непустых множеств X_1, X_2, \dots, X_p , где

$$X_j = X \cap \left(\bigcup_{i \in I_j} V G_i \right)$$

для каждого $j \in \{1, 2, \dots, p\}$. Таким образом, каждое множество X_j состоит из всех вершин доли X , принадлежащих компонентам связности $\{G_i; i \in I_j\}$. Так как каждая компонента связности G_i представляет собой полный двудольный граф, в обеих долях которого содержится по a_i вершин, то

$$|X_j| \cdot |N(X_j)| = \sum_{i \in I_j} a_i = b$$

для каждого $j \in \{1, 2, \dots, p\}$. Величина рассматриваемого разбиения множества X равна

$$\sum_{j=1}^p |X_j| \cdot |N(X_j)| = pb^2 = k.$$

Из существования разбиения X на p непустых множеств величины k немедленно следует неравенство $\zeta_p(G) \leq k$.

Обратно, пусть для графа G выполняется $\zeta_p(G) \leq k$. Это означает, что существует разбиение доли X графа G на p непустых множеств

$$X = X_1 \cup X_2 \cup \dots \cup X_p, \quad (5)$$

величина которого не больше чем k , т. е.

$$\sum_{j=1}^p |X_j| \cdot |N(X_j)| \leq k. \quad (6)$$

Если среди компонент связности графа G есть хотя бы одна компонента G_t , $t \in \{1, 2, \dots, 4p\}$, такая, что вершины из множества $V G_t \cap X$ содержатся в нескольких различных элементах разбиения (5), то преобразуем это разбиение так, что величина его не возрастет и все вершины из множества $V G_t \cap X$ будут принадлежать только одному элементу разбиения.

Пусть G_t – компонента связности графа G такая, что вершины из множества $V G_t \cap X$ содержатся в множествах X_1, X_2, \dots, X_q , где $2 \leq q \leq p$, и не содержатся в множествах $X_{q+1}, X_{q+2}, \dots, X_p$. Всюду далее предполагаем, что $t > p$ (если это не так, то всегда можно перенумеровать компоненты связности G_1, G_2, \dots, G_{4p} соответствующим образом). Пусть s – число множеств X_1, X_2, \dots, X_q , содержащих только вершины из множества $V G_t \cap X$. Тогда возможны только следующие два случая.

С л у ч а й 1. Пусть $s = 0$. Среди множеств X_1, \dots, X_q найдем множество, скажем X_{q^*} , окружение которого состоит из наименьшего числа вершин. Сейчас переместим из каждого множества X_i , $i \in \{1, 2, \dots, q\} \setminus \{q^*\}$, все вершины, принадлежащие множеству $V G_t \cap X$, в множество X_{q^*} . Нетрудно видеть, что в результате такого преобразования множества X_1, \dots, X_p по-прежнему будут образовывать разбиение множества X на p непустых множеств и $V G_t \cap X \subset X_{q^*}$. Покажем, что такое преобразование не увеличит величину разбиения. Рассмотрим разбиение (5) до преобразования, произвольное множество X_i , где $i \in \{1, 2, \dots, q\} \setminus \{q^*\}$, и любую вершину $v \in X_i \cap V G_t$. Пусть для определенности $I < q^*$. Величина исходного (до преобразования) разбиения равна

$$\left(|X_1| \cdot |N(X_1)| + \dots + |X_i| \cdot |N(X_i)| + \dots + |X_{q^*}| \cdot |N(X_{q^*})| + \dots + |X_q| \cdot |N(X_q)| \right) + \sum_{\ell=q+1}^p |X_\ell| \cdot |N(X_\ell)|.$$

При перемещении вершины v из X_i в X_{q^*} в этой сумме изменятся только два слагаемых, а именно слагаемое $|X_i| \cdot |N(X_i)|$ уменьшится и станет таким $(|X_i| - 1) \cdot |N(X_i)|$ или таким $(|X_i| - 1) \cdot (|N(X_i)| - a_t)$ (в случае когда вершина v является единственной вершиной из множества $V G_t \cap X$, принадлежащей множеству X_i), а слагаемое $|X_{q^*}| \cdot |N(X_{q^*})|$ увеличится и станет таким $(|X_{q^*}| + 1) \cdot |N(X_{q^*})|$. Так как $|N(X_{q^*})| \leq |N(X_i)|$, то при перемещении вершины v из X_i в X_{q^*} сумма не увеличится, а значит и величина разбиения не увеличится. Так как величина

разбиения не увеличивается при перемещении одной (любой) вершины $v \in X_i \cap VG_t$, где $i \in \{1, 2, \dots, q\} \setminus \{q^*\}$, то величина разбиения не увеличится и при перемещении всех таких вершин (поочередно друг за другом).

С л у ч а й 2. Пусть $s > 0$. Для определенности будем предполагать, что $X_i \subset VG_t \cap X$ для каждого $i \in \{1, 2, \dots, s\}$ и $X_i \setminus (VG_t \cap X) \neq \emptyset$ для каждого $i \in \{s+1, s+2, \dots, q\}$. Рассмотрим подмножества X'_1, X'_2, \dots, X'_p множества X , где

$$X'_i = \begin{cases} VG_t \cap X, & \text{если } i = 1, \\ X_i \setminus (VG_t \cap X), & \text{если } i \in \{2, 3, \dots, q\}, \\ X_i, & \text{если } i \in \{q+1, q+2, \dots, p\}. \end{cases}$$

При этом возможно, что некоторые из множеств X'_2, X'_3, \dots, X'_q являются пустыми.

П о д с л у ч а й 2.1. Если $s = 1$, то множества X'_1, X'_2, \dots, X'_p непусты и образуют разбиение множества X на p непустых множеств и $VG_t \cap X = X'_1$. Покажем, что величина нового разбиения $X = X'_1 \cup X'_2 \cup \dots \cup X'_p$ меньше чем величина исходного разбиения $X = X_1 \cup X_2 \cup \dots \cup X_p$. Мощности множеств и их окружений исходного разбиения и нового разбиения связаны следующими соотношениями:

$$|X'_1| = |X_1| + \sum_{i=2}^q (|X_i| - |X'_i|), \quad (7)$$

$$|X'_i| < |X_i|, \quad \forall i \in \{2, 3, \dots, q\}, \quad (8)$$

$$|N(X_i)| = \begin{cases} |N(X'_i)|, & \text{если } i = 1, \\ |N(X'_i)| + |N(X'_1)|, & \text{если } i \in \{2, 3, \dots, q\}, \\ |N(X'_i)|, & \text{если } i \in \{q+1, q+2, \dots, p\}. \end{cases} \quad (9)$$

Из (7)–(9) следует, что величина разбиения $X = X'_1 \cup X'_2 \cup \dots \cup X'_p$ равна

$$\begin{aligned} & \sum_{i=1}^q |X'_i| \cdot |N(X'_i)| + \sum_{i=q+1}^p |X_i| \cdot |N(X_i)| = \left(|X_1| + \sum_{i=2}^q (|X_i| - |X'_i|) \right) \cdot |N(X'_1)| + \sum_{i=2}^q |X'_i| \cdot |N(X'_i)| + \\ & \sum_{i=q+1}^p |X_i| \cdot |N(X_i)| < \left(|X_1| + \sum_{i=2}^q (|X_i| - |X'_i|) \right) \cdot |N(X'_1)| + \sum_{i=2}^q |X_i| \cdot |N(X'_i)| + \sum_{i=q+1}^p |X_i| \cdot |N(X_i)| = \\ & \sum_{i=1}^q |X_i| \cdot |N(X_i)| - \sum_{i=2}^q |X'_i| \cdot |N(X'_1)| + \sum_{i=q+1}^p |X_i| \cdot |N(X_i)| \leq \sum_{i=1}^p |X_i| \cdot |N(X_i)|. \end{aligned}$$

Тем самым установлено, что величина нового разбиения меньше чем величина исходного разбиения.

П о д с л у ч а й 2.2. Пусть $s \geq 2$. Тогда среди множеств X'_1, X'_2, \dots, X'_q ровно $(s-1)$ пусты. Пустыми множествами являются множества X'_2, X'_3, \dots, X'_s . Так как граф G состоит из $4p$ компонент связности, то обязательно найдутся p компонент связности графа G , среди которых нет компоненты G_p , такие, что в любом множестве X'_i , $i \in \{s+1, s+2, \dots, p\}$, найдется хотя бы одна вершина, которая не содержится ни в одной из этих p компонент. Пусть для определенности эти p компонент связности графа G – это компоненты G_1, G_2, \dots, G_p (среди этих компонент связности нет компоненты связности G_p , поскольку $t > p$). Таким образом, $X'_i \setminus \left(\bigcup_{j=1}^p VG_j \cap X \right) \neq \emptyset$ для любого $i \in \{s+1, s+2, \dots, p\}$. Рассмотрим подмножества $X''_1, X''_2, \dots, X''_p$ множества X , где

$$X''_i = \begin{cases} X'_1, & \text{если } i = 1, \\ VG_{i-1} \cap X, & \text{если } i \in \{2, 3, \dots, s\}, \\ X'_i \setminus \left(\bigcup_{j=1}^{s-1} VG_j \cap X \right), & \text{если } i \in \{s+1, s+2, \dots, p\}. \end{cases}$$

Нетрудно видеть, что множества $X_1'', X_2'', \dots, X_p''$ образуют разбиение множества X на p непустых множеств и $VG_t \cap X = X_1''$. Покажем, что величина разбиения $X = X_1'' \cup X_2'' \cup \dots \cup X_p''$ не больше чем величина разбиения $X = X_1 \cup X_2 \cup \dots \cup X_p$. Тот факт, что величина разбиения $X = X_1' \cup X_2' \cup \dots \cup X_p'$ не больше чем величина разбиения $X = X_1 \cup X_2 \cup \dots \cup X_p$ доказан в подслучае 2.1. Величина разбиения $X = X_1'' \cup X_2'' \cup \dots \cup X_p''$ не больше чем величина разбиения $X = X_1' \cup X_2' \cup \dots \cup X_p'$, так как стоимость разбиения $X = X_1'' \cup X_2'' \cup \dots \cup X_p''$ не больше чем стоимость разбиения $X = X_1' \cup X_2' \cup \dots \cup X_p'$.

Итеративно применяя последовательно друг за другом описанное преобразование разбиения X_1, X_2, \dots, X_p множества X для каждой компоненты связности G_t графа G такой, что вершины $VG_t \cap X$ содержатся в нескольких различных множествах X_1, X_2, \dots, X_p , получим разбиение X_1', X_2', \dots, X_p' множества X на p непустых множеств, величина которого не больше величины исходного разбиения X_1, \dots, X_p , и все вершины каждой компоненты связности графа G , принадлежащие множеству X , содержатся ровно в одном из множеств X_1', X_2', \dots, X_p' , т. е. для любого $t \in \{1, 2, \dots, 4p\}$ существует $i \in \{1, 2, \dots, p\}$ такое, что $VG_t \cap X \subseteq X_i'$.

Рассмотрим подмножества I_1, I_2, \dots, I_p множества $I = \{1, 2, \dots, 4p\}$, где

$$I_j = \{i \in \{1, 2, \dots, 4p\} : VG_i \cap X \subseteq X_j'\}$$

для каждого $j \in \{1, 2, \dots, p\}$. Из того, что X_1', X_2', \dots, X_p' – непустые множества и при этом *все* вершины любой компоненты связности графа G , принадлежащие множеству X , содержатся ровно в одном из этих множеств следует, что множества I_1, I_2, \dots, I_p непустые и образуют разбиение множества I . Покажем, что для каждого множества I_j , $j \in \{1, 2, \dots, p\}$, выполняется равенство (3). Для этого сделаем ряд наблюдений. Заметим, что для каждого $j \in \{1, 2, \dots, p\}$

$$|X_j'| = |N(X_j')| = \sum_{i \in I_j} a_i. \quad (10)$$

Это следует из свойств множеств X_1', X_2', \dots, X_p' . Учитывая равенство (10), неравенство (6) и тот факт, что величина разбиения $X = X_1' \cup X_2' \cup \dots \cup X_p'$ не больше чем величина разбиения $X = X_1 \cup X_2 \cup \dots \cup X_p$, получаем

$$\sum_{j=1}^p \left(\sum_{i \in I_j} a_i \right)^2 = \sum_{j=1}^p |X_j'| \cdot |N(X_j')| \leq \sum_{j=1}^p |X_j| \cdot |N(X_j)| \leq k = pb^2. \quad (11)$$

Пусть теперь

$$\sum_{i \in I_j} a_i = b + \varepsilon_j, \quad \forall j \in \{1, 2, \dots, p\},$$

где ε_j – это целое число, которое показывает то насколько сумма $\sum_{i \in I_j} a_i$ отличается от числа b . Учитывая (2), получаем

$$pb = \sum_{i=1}^{4p} a_i = \sum_{j=1}^p \left(\sum_{i \in I_j} a_i \right) = \sum_{j=1}^p (b + \varepsilon_j) = pb + \sum_{j=1}^p \varepsilon_j.$$

Следовательно, $\varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_p = 0$. Учитывая (11), получаем

$$pb^2 \geq \sum_{j=1}^p \left(\sum_{i \in I_j} a_i \right)^2 = \sum_{j=1}^p (b + \varepsilon_j)^2 = pb^2 + \underbrace{2b \sum_{j=1}^p \varepsilon_j}_0 + \sum_{j=1}^p \varepsilon_j^2.$$

Откуда немедленно следует, что $\varepsilon_1 = \varepsilon_2 = \dots = \varepsilon_p = 0$. Тем самым показано, что для каждого множества I_j , $j = 1, p$, выполняется равенство (3). Теорема доказана.

Учитывая равенство (1), получаем

С л е д с т в и е. Задача «Пополнение до разбиения на биклики» в классе P_4 -свободных двудольных графов является NP-полной.

Задача в классе $2K_2$ -свободных двудольных графов. Напомним, что граф $2K_2$ представляет собой дизъюнктивное объединение двух графов изоморфных K_2 , где K_2 – полный граф на двух вершинах. Класс $2K_2$ -свободных двудольных графов является пересечением двудольных графов и графов, которые не содержат порожденных подграфов, изоморфных графу $2K_2$. Известно (см., напр., [10; 11]), что $2K_2$ -свободные двудольные графы – это в точности двудольные графы, вершины любой доли которых могут быть упорядочены таким образом, что окружение каждой вершины v упорядоченной доли содержится в окружении вершины, которая предшествует вершине v , если такая имеется. Точнее, двудольный граф $G = (X \cup Y, E)$ не содержит порожденных подграфов, изоморфных графу $2K_2$, тогда и только тогда, когда существует биективная функция $\pi: \{1, 2, \dots, |X|\} \rightarrow X$ такая, что

$$N(\pi(1)) \supseteq N(\pi(2)) \supseteq \dots \supseteq N(\pi(|X|)). \quad (12)$$

Здесь функция π реализует упорядочение вершин доли X с указанным выше свойством. Отметим, что существование такой функции π влечет наличие биективной функции, которая реализует упорядочение вершин доли Y с тем же свойством. Существует линейный алгоритм распознавания $2K_2$ -свободных двудольных графов [11]. Более того, этот алгоритм в случае, если заданный граф не содержит порожденных подграфов, изоморфных графу $2K_2$, дополнительно возвращает упорядочение π вершин доли X , которое удовлетворяет условию (12).

Л е м м а. Пусть $G = (X \cup Y, E)$ – непустой двудольный граф, для которого существует биективная функция $\pi: \{1, 2, \dots, |X|\} \rightarrow X$ такая, что выполняется (12). Тогда существует наименьшее разбиение (разбиение с наименьшей стоимостью) множества X на p непустых множеств такое, что $X_i = \{\pi(1), \pi(2), \dots, \pi(s)\}$ для некоторых $i \in \{1, 2, \dots, p\}$ и натурального числа s , удовлетворяющего неравенству $1 \leq s \leq |X| - p + 1$.

Д о к а з а т е л ь с т в о. Если в двудольном графе G есть изолированные вершины, принадлежащие доле Y , удалим их. Очевидно, это никак не повлияет на стоимость любого разбиения множества X на p непустых множеств. Отсутствие изолированных вершин в доле Y гарантирует, что $N(\pi(1)) = Y$.

Рассмотрим наименьшее разбиение множества X на p непустых множеств X_1, X_2, \dots, X_p . Не теряя общности, допустим, что $\pi(1) \in X_1$. Если $X_1 = \{\pi(1), \pi(2), \dots, \pi(|X_1|)\}$, то лемма доказана. В противном случае существуют два натуральных числа a, b такие, что $1 < a < b \leq |X|$ и $\pi(a) \notin X_1$, $\pi(b) \in X_1$. Пусть $\pi(a) \in X_j$, $j \neq 1$. Преобразуем рассматриваемое разбиение, поменяв местами вершины $\pi(a)$ и $\pi(b)$. В результате получим разбиение множества X на p непустых множеств X'_1, X'_2, \dots, X'_p , где

$$X'_i := \begin{cases} (X_i \setminus \{\pi(b)\}) \cup \{\pi(a)\}, & \text{если } i = 1, \\ (X_i \setminus \{\pi(a)\}) \cup \{\pi(b)\}, & \text{если } i = j, \\ X_i, & \text{если } i \neq 1, i \neq j. \end{cases}$$

Утверждается, что стоимость получившегося разбиения не больше чем стоимость разбиения $X = X_1 \cup X_2 \cup \dots \cup X_p$. Для того чтобы это доказать, достаточно показать, что

$$c(G[X'_1 \cup N(X'_1)]) + c(G[X'_j \cup N(X'_j)]) \leq c(G[X_1 \cup N(X_1)]) + c(G[X_j \cup N(X_j)]). \quad (13)$$

Так как $\pi(1) \in X_1$ и $\pi(1) \in X'_1$, то $N(X_1) = N(X'_1) = Y$. Учитывая это и тот факт, что $N(\pi(b)) \subseteq N(\pi(a))$, имеем

$$c(G[X'_1 \cup N(X'_1)]) = c(G[X_1 \cup N(X_1)]) - |N(\pi(a)) \setminus N(\pi(b))|. \quad (14)$$

Так как $N(\pi(b)) \subseteq N(\pi(a))$, то $N(X'_j) \subseteq N(X_j)$ и

$$c(G[X'_j \cup N(X'_j)]) \leq c(G[X_j \cup N(X_j)]) + |N(\pi(a)) \setminus N(\pi(b))|. \quad (15)$$

Из соотношений (14) и (15) немедленно следует неравенство (13).

Осуществив преобразования разбиения X_1, X_2, \dots, X_p множества X такого типа последовательно друг за другом для каждой пары натуральных чисел a, b таких, что $1 < a < b \leq |X|$ и $\pi(a) \notin X_1, \pi(b) \in X_1$, получим разбиение наименьшей стоимости множества X на p непустых множеств, из которых одно состоит из вершин $\pi(1), \pi(2), \dots, \pi(s)$ для некоторого натурального числа s такого, что $1 \leq s \leq |X| - p + 1$. Лемма доказана.

Используя математическую индукцию по числу множеств, на которые разбивается множество X , можно показать, что для любого $2K_2$ -свободного двудольного графа $G = (X \cup Y, E)$ существует наименьшее разбиение множества X на p непустых множеств X_1, X_2, \dots, X_p следующего вида

$$\begin{aligned} X_1 &= \{\pi(1), \pi(2), \dots, \pi(s_1)\}, \\ X_2 &= \{\pi(s_1 + 1), \pi(s_1 + 2), \dots, \pi(s_2)\}, \\ &\dots \\ X_p &= \{\pi(s_{p-1} + 1), \pi(s_{p-1} + 2), \dots, \pi(|X|)\}, \end{aligned} \tag{16}$$

где $1 \leq s_1 < s_2 < \dots < s_{p-1} < |X|$. Следовательно, для того чтобы найти разбиение наименьшей стоимости множества X на p непустых множеств достаточно осуществить перебор $(p - 1)$ -элементных последовательностей из множества

$$\{(s_1, s_2, \dots, s_{p-1}) : 1 \leq s_1 < s_2 < \dots < s_{p-1} < |X|\}.$$

Существует алгоритм, который находит наименьшее разбиение доли X на p непустых множеств вида (16) за время $O(p|X|^4|Y|)$. Для любой четверки натуральных чисел (i, j, ℓ, q) такой, что $1 \leq i \leq \ell \leq j \leq |X|$, $1 \leq q \leq p$, $j - i + 1 \geq q$, $j - \ell \geq q - 1$, через $f(i, j, \ell, q)$ обозначим наименьшую стоимость разбиения множества вершин $\{\pi(i), \pi(i+1), \dots, \pi(j)\}$ на q непустых множеств, среди которых, при условии что $q \geq 2$, обязательно имеется множество, состоящее только из вершин $\pi(i), \pi(i+1), \dots, \pi(\ell)$.

Алгоритм последовательно вычисляет значения всех величин $f(i, j, \ell, q)$ в порядке роста значений q , используя следующее рекуррентное соотношение: если $q = 1$, то для любого ℓ значение $f(i, j, \ell, q)$ равно недостатку подграфа графа G , который порожден множеством вершин $\{\pi(i), \pi(i+1), \dots, \pi(j)\} \cup N(\{\pi(i), \pi(i+1), \dots, \pi(j)\})$, причем этот недостаток может быть найден за время $O(|X||Y|)$; если $q \geq 2$, то

$$f(i, j, \ell, q) = \sum_{t=i}^{\ell} (|N(\pi(t))| - |N(\pi(t))|) + \min_{\ell'} f(\ell + 1, j, \ell', q - 1),$$

где минимум берется по всем натуральным числам ℓ' таким, что $\ell + 1 \leq \ell' \leq j - (q - 2)$. В этой формуле первое слагаемое – недостаток подграфа графа G , который порожден множеством вершин $\{\pi(i), \pi(i+1), \dots, \pi(\ell)\} \cup N(\{\pi(i), \pi(i+1), \dots, \pi(\ell)\})$, второе слагаемое – наименьшая стоимость разбиения множества вершин $\{\pi(\ell + 1), \pi(\ell + 2), \dots, \pi(j)\}$ на $q - 1$ непустых множеств, среди которых, при условии $q \geq 3$, существует множество, состоящее из вершин $\pi(\ell + 1), \pi(\ell + 2), \dots, \pi(\ell')$ по всем ℓ' таким, что $\ell + 1 \leq \ell' \leq j - (q - 2)$. Наименьшая стоимость разбиения X на p непустых множеств может быть найдена так:

$$\xi_p(G) = \min_{\ell} f(1, |X|, \ell, p),$$

где минимум берется по всем натуральным числам ℓ , не превосходящим $|X| - p + 1$. Оценим трудоемкость этого алгоритма в предположении, что граф представлен в виде списков смежности. В ходе работы алгоритма значения $f(i, j, \ell, q)$ вычисляются для четверок (i, j, ℓ, q) , которых не более чем $p|X|^3$. Вычисление значения $f(i, j, \ell, q)$ для одной четверки (i, j, ℓ, q) потребует $O(|X||Y|)$ времени. Следовательно, трудоемкость алгоритма составляет $O(p|X|^4|Y|)$.

Заключение. В работе показано, что задача «Пополнение до разбиения на биклики» является NP -полной в классе P_4 -свободных двудольных графов. Тем самым установлено, что эта задача NP -полна в таких широко изучаемых надклассах P_4 -свободных двудольных графов, как двудольные перестановочные графы, двудольные хордалльные графы, двудольные дистанционно-наследственные графы, P_5 -свободные двудольные графы, $S_{1,2,3}$ -свободные двудольные графы. Также предлагается алгоритм для решения этой задачи в классе $2K_2$ -свободных двудольных графов.

Авторы благодарят В. В. Лепина за обращение их внимания на рассматриваемую в работе задачу и за полезные замечания, которые способствовали значительному улучшению работы. Первый автор был поддержан Белорусским республиканским фондом фундаментальных исследований (проект № Ф14РА-004).

Список использованной литературы

1. Лекции по теории графов / В. А. Емеличев [и др.]. – М.: Наука, 1990.
2. Biclique completion problems for multicast network design / N. Faure [et al.] // *Discrete Optimization*. – 2007. – Vol. 4. – P. 360–377.
3. *Faure, N.* Contribution à la resolution de problèmes de regroupement de sessions multicasts: PhD thesis / N. Faure. – Université Paris VI, 2006 (in French).
4. *Gualandi, S.* Proceedings of the 6th International Conference on Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems / S. Gualandi. – Pittsburgh, USA: Springer Berlin Heidelberg, 2009. – P. 87–101.
5. *Gualandi, S.* A branch-and-price approach to k-clustering minimum biclique completion problem / S. Gualandi, F. Maffioli, C. Magni // *International transactions in operational research*. – 2013. – Vol. 20. – P. 101–117.
6. *Magni, C.* Biclique completion problem: models and algorithms: MSc thesis / C. Magni. – Politecnico di Milano, 2009.
7. *Gualandi, S.* Weighted Biclique Completion via CP-SDP Randomized Rounding / S. Gualandi, F. Malucelli // *Proceedings of the European Workshop on Mixed Integer Nonlinear Programming*. – Marseille, France, 2010. P. 223–230.
8. *Гэри, М.* Вычислительные машины и труднорешаемые задачи / М. Гэри, Д. Джонсон. – М.: Мир, 1982.
9. *Babel, L.* Recognizing the P4-structure of bipartite graphs / L. Babel, A. Brandstädt, V. B. Le // *Discrete Applied Mathematics*. – 1999. – Vol. 93. – P. 157–168.
10. *Feder, T.* Approximating the minimum chain completion problem / T. Feder, H. Mannila, E. Terzi // *Information Processing Letters*. – 2009. – Vol. 109. – P. 980–985.
11. *Heggernes, P.* Linear-time certifying recognition algorithms and forbidden induced subgraphs / P. Heggernes, D. Kratsch // *Nordic J. of Computing*. – 2007. – Vol. 14. – P. 87–108.

Поступило в редакцию 08.07.2015