PCPs and Inapproximability CIS 6930 November 5, 2009

Lecture Balanced Connected Partitions

Lecturer: Dr. My T. Thai Scribe: Yilin Shen

1 Problem Definition

Definition 1 (Max Balanced Connected q-Partition (BCP_q)) *Given a connected graph $G = (V, E)$ with a weight function $w : V \to \mathbb{Z}_+$ and $q \geq 2$ be a positive integer. For $X \subseteq V$, let $w(X)$ denote the sum of the weights of the vertices in X. The BCP$_q$ problem on G is to find a q-partition $P = (V_1, V_2, \ldots, V_q)$ of V such that $G[V_i]$ is connected $(1 \leq i \leq q)$ and P maximizes $\min\{w(V_i) : 1 \leq i \leq q\}$.*

Definition 2 (Exact Cover by 3-Sets (X3C)) *Given a set X with $|X| = 3q$ and a family C of 3-element subsets of X, $|C| = 3q$, where each element of X appears in exactly 3 sets of C, decide whether C contains an exact cover for X, that is, a subcollection $C' \subseteq C$ such that each element of X occurs in exactly one member of C'.*

2 Related works

- The simpler unweighted version of BCP_q is the special case of BCP_q in which all vertices have weight 1 (denoted by 1-BCP_q):
  - For every $q \geq 2$, the problem 1-BCP_q is *NP*-hard (even for bipartite graphs);
  - When the input graph has a higher connectivity, we have: Let G be a q-connected graph with n vertices, $q \geq 2$, and let $n_1, n_2, \ldots, n_q$ be positive natural numbers such that $n_1 + n_2 + \ldots + n_q = n$. Then G has a connected q-partition $(V_1, V_2, \ldots, V_q)$ such that $|V_i| = n_i$ for $i = 1, 2, \ldots, q$.
- The more general weighted case:
  - BCP_q is polynomially solvable only for ladders and for trees;
  - BCP_q restricted to grids $G_{m\times n}$ with $n \geq 3$ is already *NP*-hard;
  - BCP_2 is *NP*-hard on connected graphs, bipartite graphs, and graphs with at least one block containing $(\log|V|)$ articulation points and complete graphs.

3 Hardness of Approximation

Theorem 3 *The decision version of BCP_2 is* NP*-complete in the strong sense for 2-connected graphs.*

Proof:

We will show the reduction from $X3C$ problem to BCP_2 problem

The Construction:

Given an instance (X, C) of $X3C$, let $G = (V, E)$ be the graph with vertex set $V = X \cup C \cup a, b$ and edge set $E = \bigcup_{j=1}^{3q} [C_j x_i | x_i \in C_j \cup C_j a \cup C_j b]$. Clearly, G can be constructed in polynomial time in the size of (X, C). It is also not difficult to see that G is 2-connected.

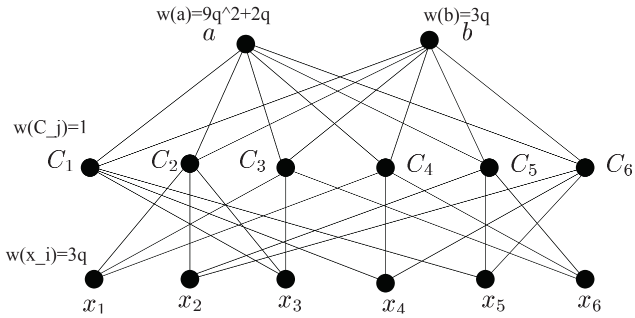

Figure 1: Example: Graph obtained by the reduction for the instance (X, C), where $C = \{C_1, C_2, \ldots, C_6\}$, $C_1 = \{x_3, x_4, x_5\}$, $C_2 = \{x_1, x_2, x_3\}$, $C_3 = \{x_1, x_3, x_6\}$, $C_4 = \{x_1, x_4, x_6\}$, $C_5 = \{x_2, x_5, x_6\}$ and $C_6 = \{x_2, x_4, x_5\}$

Define a weight function $w : V \rightarrow \mathbb{Z}_+$ as follows: $w(a) = 9q^2 + 2q$; $w(b) = 3q$; $w(x_i) = 3q$ for $i = 1, \ldots, 3q$; and $w(C_j) = 1$ for $j = 1, \ldots, 3q$. Note that $w(V) = 2(9q^2 + 4q)$.

We will prove that C contains an exact cover for X if and only if G has a connected 2-partition (V_1, V_2) such that $\min\{w(V_1), w(V_2)\} \geq W/2$, where $W = w(V)$:

($\Rightarrow$) Given an exact cover C', consider the connected 2-partition (V_1, V_2) of G, where $V_1 = \{a\} \cup \{C_j : C_j \notin C'\}$ and $V_2 = \{b\} \cup \{C_j, x_i : C_j \in C', x_i \in C_j\}$. Since C' consists of q subsets, we have $w(V_1) = w(a) + 3q \cdot q = 9q^2 + 4q = W/2$.

($\Leftarrow$) Let (V_1, V_2) be a connected 2-partition of G such that $\min\{w(V_1), w(V_2)\} \geq W/2$, *w.l.o.g*, that is, $\min\{w(V_1), w(V_2)\} = W/2$. *Note that a and b cannot belong to the same set V_i, because $w(a) + w(b) = 9q^2 + 5q > W/2$.*

Suppose that $a \in V_1$ and $b \in V2$, we know no vertex of X is in V_1, otherwise $w(V_1) \geq w(a) + 3q > W/2$. Therefore V_1 contains the vertex a and some vertices of C. Since $w(V_1) = W/2 = 9q^2 + 4q$, this implies that V_1 contains exactly $2q$ vertices of C. Thus, V_2 has precisely q vertices of C. Since these q vertices are independent and the vertices in $X \cup b$ are also independent, it is easy to verify that these q vertices of C belonging to V_2 define an exact cover for X since V_2 is connected.

The proof is complete.

■

Theorem 4 *There is no* $(1+\epsilon)$*-approximation algorithm for the problem* BCP_2*, where* $\epsilon \le 1/n^2$ *and* n *is the number of vertices of the input graph, unless* P = NP.

Proof:

We will show the *gap-producing reduction* from $X3C$ problem to BCP_2 problem

The Construction is similar as Theorem 3 except the weight assignment as follows (Shown in Figure 2.): $w(a) = 6q^3 + q^2$; $w(b) = 2q^2$; $w(C_j) = q$ for $j = 1, \dots, 3q$; and $w(x_i) = 2q^2$ for $i = 1, \dots, 3q$. Observe that $w(V) = 2(6q^3 + 3q^2)$ and $|V| = 6q + 2$.

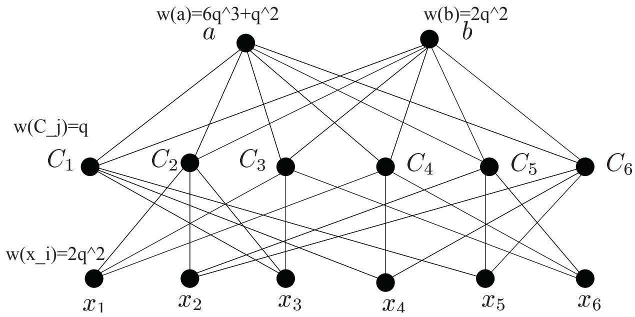

Figure 2: Example: Graph obtained by the reduction for the instance (X, C), where $C = \{C_1, C_2, \dots, C_6\}$, $C_1 = \{x_3, x_4, x_5\}$, $C_2 = \{x_1, x_2, x_3\}$, $C_3 = \{x_1, x_3, x_6\}$, $C_4 = \{x_1, x_4, x_6\}$, $C_5 = \{x_2, x_5, x_6\}$ and $C_6 = \{x_2, x_4, x_5\}$

Then we will show the completeness and soundness:

- **Completeness:** $\exists$ Exact Cover $C' \Rightarrow w(V_1) = w(V_2) = w(V)/2$

  Given an exact cover C' of X, consider the connected 2-partition (V_1, V_2) of G, where $V_1 = \{a\} \cup (C \setminus C')$ and $V_2 = \{b\} \cup X \cup C'$. Clearly, we have that $w(V_1) = 6q^3 + 3q^2 = w(V)/2 = w(V_2)$.

- **Soundness:** $\nexists$ Exact Cover $C' \Rightarrow \min\{w(V_1), w(V_2)\} < w(V)/2 - q + 1$

  Assume that (V_1, V_2) be a connected 2-partition of G with measure m such that $m \ge w(V)/2 - q + 1$, we will prove the contradiction. There are two scenarios:

  - a and b belong to the same set V_i $(i = 1, 2)$: In this case, a, b and at least one C_j for some $j \in \{1, 2, ..., 3q\}$ would belong to the same set in order to induce a connected partition. Therefore, the weight of this set would be at least $6q^3 + 3q^2 + q = w(V)/2 + q$, that is, $m \le w(V)/2 - q$, a contradiction.
  - Suppose that $a \in V_1$ and $b \in V_2$.

    We will prove that $X \in V_2$ first. Suppose there exists at least one x_i in V_1, at least one C_j has to be selected in V_1 to make the partition connected. Therefore, $w(V_1) \ge 6q^3 + q^2 + 2q^2 + q = w(V)/2 + q$, that is, $m \le w(V)/2 - q$, a contradiction.

    Since $X \in V_2$ and the subgraph induced by V_2 has to be connected, we have $|C \cap V_2| \ge q$. If $|C \cap V_2| > q$, which implies that $|C \cap V_1| \le 2q - 1$. In this case, $w(V_1) \le 6q^3 + q^2 + (2q-1)q = 6q^3 + 3q^2 - q = w(V)/2 - q$, and we have a contradiction again. Thus $|C \cap V_2| = q$, and hence $C \cap V_2$ covers X exactly.

And since $n = 6q+2$ and $w(V) = 2(6q^3+3q^2)$, it's easily known that $w(V)/2(n^2+1) < q$. Therefore, we know that BCP_2 cannot be approximated within ratio

$$\frac{w(V)/2}{w(V)/2 - q + 1} = \frac{1}{1 - \frac{q-1}{w(V)/2}} > \frac{1}{1 - \frac{q-1}{q(n^2+1)}} = \frac{q(n^2+1)}{qn^2 - 1} > \frac{n^2+1}{n^2} > 1 + \epsilon$$

where $\epsilon < 1/n^2$. Since there are no $\frac{w(V)/2}{w(V)/2-q+1} - \varepsilon$ approximation algorithm and according to the above equation, there is no $(1+\epsilon)$-approximation algorithm for the problem BCP_2, where $\epsilon \leq 1/n^2$.

The proof is complete.

■

Theorem 5 *For every $q \geq 2$, the decision version of BCP_q is* NP*-complete in the strong sense for q-connected graphs.*

Proof:

Denote by $DBCP_q$ the decision version of BCP_q. Suppose $q \geq 3$. We prove, by induction on q, that the problem $DBCP_q - 1$ can be reduced to the problem $DBCP_q$.

The Construction:

Let $I = (G, w, m)$ be an instance of $DBCP_{q-1}$ that consists of a $(q-1)$-connected graph $G = (V, E)$, a function $w : V \to \mathbb{Z}_+$ and a positive integer m. The goal is to decide whether this instance has a solution with measure at least m.

We construct an instance $I' = (G', w', m)$ of $DBCP_q$ that consists of a q-connected graph $G' = (V', E')$, with $V' = V \cup \{v'\}$, where $v' \notin V$, and $E' = E \cup \{v'u : u \in V\}$, and a function w' on the vertices of G' such that: $w'(v') = w(V)/(q-1)$ and $w'(v) = w(v)$ for each $v \in V$. It is obvious that G' can be constructed in polynomial time in the size of I and G' is q-connected.

We will prove that the instance I of $DBCP_{q-1}$ has a connected $(q-1)$-partition with measure at least m if only if the instance I' of $DBCP_q$ has a connected q-partition with measure at least m.

($\Rightarrow$) Let $P = (X_1, \ldots, X_{q-1})$ be a connected $(q-1)$-partition of G with measure at least m. In this case, $(X_1, \ldots, X_{q-1}, \{v'\})$ is a connected q-partition of G' with measure at least m.

($\Leftarrow$) Suppose that $P' = (X'_1, \ldots, X'_q)$ is a connected q-partition of G' with measure m', where $m' \geq m$. *w.l.o.g*, suppose that X'_q contains v' and $w'(X'_1) \geq w'(X'_i)$ for $2 \leq i \leq q-1$. Since $w'(V') = w(V) + w(V)/(q-1)$ and $w'(X'_q) \geq w(V)/(q-1)$, we have $w'(X'_q) \geq w'(V')/q$. Thus, we know that $m' = w'(X'_1))$. Let $R = X'_q \setminus \{v'\}$, we have two scenarios:

- If $R = \emptyset$, $P' = (X'_1, \ldots, X'_{q-1})$ is a connected $(q-1)$-partition of G with measure m'.
- If $R \neq \emptyset$, since G is $(q-1)$-connected (weaker than q-connected), there exists a way to distribute the vertices of R among the sets X'_i for $1 \leq i \leq q-1$ in such a way that the new sets $X'_i \cup R_i$, where $\cup_{i=1}^{q-1} R_i = R$, which induce the connected subgraphs of G. In this case, $(X'_1 \cup R_1, \ldots, X'_{q-1} \cup R_{q-1})$ is a connected $(q-1)$-partition of G with measure at least m'. Since $m' \geq m$, the proof of the claim is complete.

■

Theorem 6 *The problem BCP does not admit an α-approximation algorithm with $\alpha < 6/5$, unless* P = NP.

Proof:

We will show the reduction from $X3C$ problem to BCP problem

The Construction:

Let $I = (X, C)$ be an instance of X3C, where $C = \{C_1, C_2, \ldots, C_{3q}\}$ is a family of subsets of $X = \{x_1, x_2, \ldots, x_{3q}\}$. Let $\epsilon > 0$ be a small number. Construct an instance $I' = I'(\epsilon) = (G = (V, E), w, Q)$ (where Q is the number of connected partitions) of BCP in the following way:

- Let $Q = 10q$.
- For each x_i in X, let $H(x_i)$ be the graph with 16 vertices, we will call gadget, defined as follows. It consists of 3 vertical paths of length 3, say P_1, P_2 and P_3, all ending at a common vertex x_i, and internally vertex-disjoint. Each such a path P_j starts at a vertex named t_{i,i_j}. The start vertices t_{i,i_1}, t_{i,i_2}, t_{i,i_3} correspond to the 3 sets C_{i1}, C_{i2}, C_{i3} that contain x_i. These vertices will be referred as *t-vertices*. For each of the 3 possible choices of two paths (among P_1, P_2 and P_3), we attach two other new vertices, as follows. Let $Pj = (t_{i,i_j}, z_{i,j}, y_{i,j}, x_i)$, for $j = 1, 2, 3$. Take two new vertices $l_{i,1}$ and $r_{i,1}$ and attach each of them to the vertices $y_{i,1}$ and $y_{i,2}$; take two other new vertices $l_{i,2}$ and $r_{i,2}$ and attach each of them to the vertices $z_{i,2}$ and $y_{i,3}$; take two other new vertices $l_{i,3}$ and $r_{i,3}$ and attach each of them to the vertices $z_{i,1}$ and $z_{i,3}$.

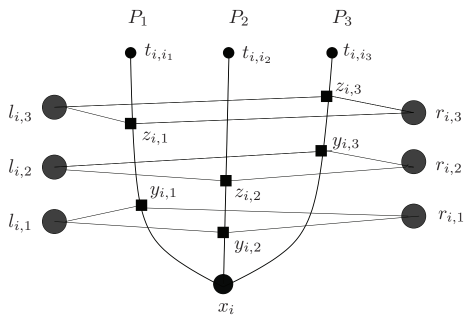

Figure 3: The gadget $H(x_i)$

- Let $G = (V, E)$ be the graph obtained from the union of the gadgets $H(x_i)$, $i = 1, \ldots, 3q$ with some additional $3q$ vertices and $9q$ edges, in the following way. Let $v_1, v_2, \ldots, v_{3q}$ be the additional vertices, where each v_i corresponds to a set C_i of the instance I of X3C. Now, whenever there is a set $C_p = \{x_i, x_j, x_k\}$ in the instance I, add three edges linking vertex v_p to the vertices $t_{i,p}$, $t_{j,p}$ and $t_{k,p}$ of the gadgets $H(x_i)$, $H(x_j)$, and $H(x_k)$, respectively. The vertices v_j will be called *v-vertices*.
- Let n be the number of vertices of the graph G (note that $n = 51q$), and let a be an integer such that $a \geq n/\epsilon$.
- The weight function $w : V \to \mathbb{Z}_+$ is defined as follows. We assign weight $2a$ to the vertices x_i; weight $3a$ to the vertices $l_{i,j}$ and $r_{i,j}$, $i = 1, \ldots, 3q$, $j = 1, 2, 3$;

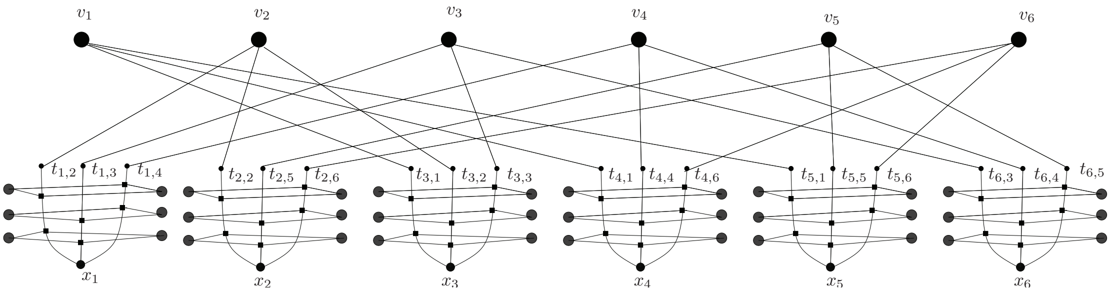

Figure 4: The gadget $H(x_i)$

and weight 1 to the remaining vertices. Note that each gadget has weight $20a+9$ and $w(V) = (60a + 30)q$.

Then we will show the completeness and soundness:

- **Completeness:** $\exists$ Exact Cover $C' \Rightarrow I'$ has a solution with measure $6a + O(1)$

  Let $C_{i_1}, C_{i_2}, \ldots, C_{i_q}$ be an exact cover for I. Construct a connected $10q$-partition for the instance I' as follows. First, construct q connected classes by considering the q sets in the exact cover. (For example, for the instance I corresponding to the graph shown in Figure 4, consider the exact cover consisting of C_3 and C_6.) Each of these connected classes consists of a vertex v_j corresponding to a set C_{i_j} together with the 3 edges leaving it, each of them extended (in a connected way) with the unique vertical path that starts at one of its extremes (a t-vertex). Clearly, each of these q connected classes has weight $6a + 10$. Consider the graph G0 obtained from G after removing the vertices in the q connected classes we have constructed so far. The other $9q$ connected classes can be obtained from G' as follows: first, in the remaining part of each gadget $H(x_i)$, construct 3 sets of paths, each one linking pairs of vertices of type $l_{i,j}$ and $r_{i,j}$. Note that this is possible, as only the vertices of one vertical path in each gadget were removed. Now, put each of the remaining vertices (all of weight 1) in any of the $10q$ connected classes constructed so far, so as to obtain a connected 10q-partition of G. Clearly, all the connected classes have weight at least $6a + 1$.

- **Soundness:** $\nexists$ Exact Cover $C' \Rightarrow I'$ has a solution with measure at most $5a + O(1)$

  Let $(V_1, V_2, \ldots, V_Q)$ be a solution of I' with measure at least $(5 + \epsilon)a$. Since $n \leq \epsilon a$, the connected classes in this solution all have to contain one or more vertices with weight $2a$ or $3a$, and therefore the weight of any connected class $G[V_j]$ must satisfy $K_j a \leq w(V_j) < K_j a + \epsilon a$, for some integer K_j.

  Since the average weight of a connected class is $6a + 3$, if there existed a connected class with weight at least $7a$, then there would exist another connected class with weight at most $5a + 6$, and therefore smaller than $(5 + \epsilon)a$, a contradiction to our hypothesis.

  Thus, $w(V_j) < 7a$, and therefore $w(V_j) = 6a + o(a)$, for $j = 1, \ldots, Q$. Thus each connected class must contain either 2 vertices with weight $3a$ or 3 vertices with weight $2a$. Let Y be a connected class containing xi. Suppose Y contains no t-vertex. Then Y can additionally contain only vertices with weights 1 or $3a$

in the gadget $H(x_i)$, and therefore it will not have weight $6a + o(a)$. Thus, Y must contain a t-vertex of the gadget $H(x_i)$. Since Y has weight $6a+o(a)$, it has to contain precisely 3 vertices with weight $2a$ and some vertices of weight 1. But the only way to connect 3 vertices with weight $2a$ is passing through a v-vertex v_j. Since the v-vertices have degree 3, one of the two cases may happen: (1) either Y contains exactly one vertex v_j, or (2) Y contains at least two v-vertices.

In case (2), Y must contain two t-vertices belonging to a same gadget, and furthermore they must be connected by a path contained in this gadget. In this case, since no vertex with weight $3a$ can be used in such a path, two vertical paths in this gadget must be used. But then, these vertical paths separate a pair of vertices with weight $3a$, and therefore some connected class will have weight $3a$, a contradiction. Thus, case (1) must occur, and in this case, Y contains 3 vertices with weight $2a$, precisely one vertical path in the corresponding 3 gadgets and one vertex v_j. Since each vertex x_i belongs to a connected class with precisely one v-vertex, there are exactly q connected classes that induce an exact cover for the instance I of X3C.

This completes the proof.
■

4 Strongly *NP*-hard Problems

Definition 7 (Strongly *NP*-hard) *A problem is strongly* NP*-hard if every problem in* NP *can be polynomially reduced to it in such a way that all of its numerical parameters in the reduced instance are always a polynomial in the length of the input.*

Definition 8 (Pseudo-polynomial algorithm) *An algorithm for a problem is a pseudo-polynomial algorithm for that problem if its running time is bounded by a polynomial.*

Theorem 9 *Consider a integral-valued strongly* NP*-hard minimization problem II with the following restriction: Let B be a numerical bound which is polynomial of a given weakly* NP*-hard problem. On any instance of II, the optimal solution is at most B.*

Then II do NOT have a FPTAS.

Proof:

We will prove that if II admits an FPTAS, then it also admits a pseudo-polynomial algorithm.

Assume II admits an FPTAS, we run FPTAS on a problem instance with $\epsilon = 1/B$. The returned value is at most the following.

$$(1+\epsilon)OPT < OPT + \epsilon B = OPT + 1$$

Since the running time is polynomial in B (which is polynomial in the size of the given weakly *NP*-hard problem), II has a pseudo-polynomial algorithm. ■

References

[1] Y. Wakabayashi, F. Chataigner, and L. B. Salgado. Approximation and inapproximability results on balanced connected partitions of graphs. *Discrete Mathematics & Theoretical Computer Science*, 9(1), 2007.

[2] P. M. Camerini, G. Galbiati, and F. Maffioli. On the complexity of finding multi-constrained spanning trees. *Discrete Appl. Math*, 5(1):39-50, 1983.