Cycle Optimization is NP-Complete

Philip Chong

January 2, 2003

1 Problem Definition

Let $G = (V, E)$ be a directed graph and let $\phi : V \to \mathbb{Z}^*$ and $\theta : E \to \mathbb{Z}^*$ be weighting functions on the vertices and edges of G, respectively, to the non-negative integers. For any cycle K in G, define the *cycle throughput* of K with respect to the weighting functions ϕ and θ to be

$$T(K, \phi, \theta) = \frac{\sum_{v \in K} \phi(v)}{\sum_{v \in K} \phi(v) + \sum_{e \in K} \theta(e)}$$

The *cycle optimization problem* (CO) is: given G, ϕ and θ, and some target throughput $q \in \mathbb{Q}$, find $\phi' : V \to \mathbb{Z}^*$ and $\theta' : E \to \mathbb{Z}^*$ such that

$$c = \sum_{v \in V} |\phi(v) - \phi'(v)| + \sum_{e \in E} |\theta(e) - \theta'(e)|$$

is minimized, subject to the condition that, for all cycles K in G,

$$T(K, \phi', \theta') \geq q$$

1.1 Notes on Definition

If a vertex is split or collapsed, $\phi'(v) > \phi(v)$ or $\phi'(v) < \phi(v)$, respectively. Likewise, if relay stations are inserted or deleted on an edge, $\theta'(e) > \theta(e)$ or $\theta'(e) < \theta(e)$, respectively.

Note that we should never consider collapsing vertices; such a move only decreases the cycle throughputs of all cycles which contain the collapsed vertex, and thus only increases the cost function without any benefit in satisfying the cycle conditions. That is, note that $T(K, \phi', \theta')$ is monotonically increasing in ϕ'. Suppose (ϕ', θ') is a solution to some instance of CO, where $\phi'(v) < \phi(v)$ for some $v \in V$. Then we can construct $\phi''(v) = \max(\phi'(v), \phi(v))$, so that (ϕ'', θ') satisfies all the conditions of the CO instance, and $\sum_{v \in V} |\phi(v) - \phi''(v)| < \sum_{v \in V} |\phi(v) - \phi'(v)|$ so (ϕ', θ') does not minimize c, a contradiction. Thus $\phi'(v) \geq \phi(v)$ for all $v \in V$. Likewise, we should never consider adding relay stations, as $T(K, \phi', \theta')$ is monotonically decreasing in θ'. Thus $\theta'(e) \leq \theta(e)$, so we may write

$$c = \sum_{v \in V} (\phi'(v) - \phi(v)) + \sum_{e \in E} (\theta(e) - \theta'(e))$$

Normally when splitting vertices, one would want to add a new vertex to the graph, and an edge with zero weight from the split vertex to the new vertex, and adjust the inedges and outedges of the split vertex correspondingly. However, since no relay stations are to be added, there is no need to make the new vertex and edge explicit, since there is no change in the weighting on the new edge possible. The weighting of a vertex is thus sufficient to reflect the number of times a vertex is split, plus 1.

1.2 Decision Problem

The cycle optimization decision problem (COD) related to CO is: given G, ϕ, θ, q, along with a target cost $c' \in \mathbb{Z}^*$, determine whether there exist $\phi' : V \to \mathbb{Z}^*$ and $\theta' : E \to \mathbb{Z}^*$ such that $c \leq c'$ and subject to the condition that, for all cycles K in G, $T(K, \phi', \theta') \geq q$ (with c and T defined as above).

2 Complexity Analysis

Theorem 1. *COD is NP-complete.*

Proof of Theorem 1. The problem lies in NP. The pair (ϕ', θ') serves as a certificate for a solution. The certificate can be checked in polynomial time to verify that it does satisfy the constraint $T(K, \phi', \theta') \geq q$ for all cycles K (this is equivalent to the Maximum Cycle Mean problem, known to be in P), and verification that $c \leq c'$ can be done in time linear in the size of the graph.

To show completeness, we reduce 3-SAT to COD. Let X be the set of variables in a 3-SAT instance, and let Y be the set of clauses from the CNF 3-SAT instance. We assume that $|Y| \geq 3$. If the SAT formula has fewer than three clauses, we can duplicate clauses from the formula until we have at least 3 clauses. This can be done in polynomial time, and does not change the solution in any way; a variable assignment satisfies the new formula iff it satisfies the original formula as well. We will utilize the property $|Y| \geq 3$ later in this proof.

For every $x \in X$, construct a set of vertices

$$V_x = \{v_{x,y}, v_{x',y} : y \in Y\}$$

(i.e. V_x contains two vertices for every clause in the 3-SAT formula). Associate $v_{x,y}$ with the positive literal x and $v_{x',y}$ with the negative literal x'. Take

$$V = \bigcup_{x \in X} V_x$$

Note that there is a one-to-one mapping from the vertices in V to the set $(X \cup X') \times Y$, where $X' = \{x' : x \in X\}$.

Now for every $x \in X$, construct a set of edges

$$E_x = \{(v_{x,y_1}, v_{x',y_2}), (v_{x',y_2}, v_{x,y_1}) : y_1, y_2 \in Y\}$$

so that $G_x = (V_x, E_x)$ is a complete bipartite graph on the vertices of V_x, where the vertices are partitioned between those corresponding to positive literals and those

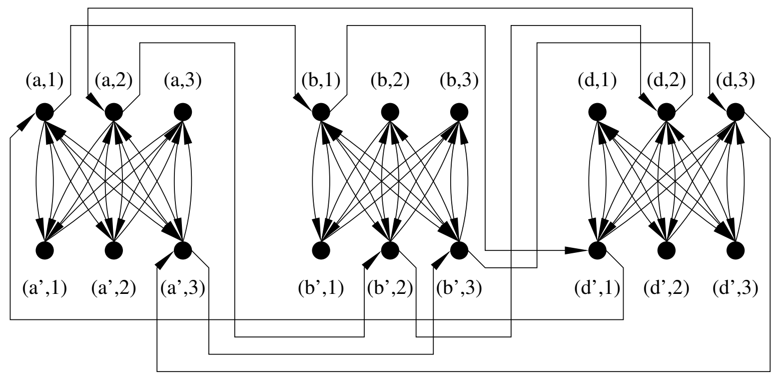

Figure 1: Example Reduction to COD for $(a + b + d')(a + b' + d)(a' + b' + d)$

corresponding to negative literals. Also, for every $y = (x_{y1} + x_{y2} + x_{y3}) \in Y$, where the x_{yi} are the literals in the CNF clause, construct a set of edges

$$E'_y = \big\{(v_{x_{y1},y}, v_{x_{y2},y}), (v_{x_{y2},y}, v_{x_{y3},y}), (v_{x_{y3},y}, v_{x_{y1},y})\big\}$$

That is, for each clause y, take the set of literals which appear in that clause, and for each literal x_{yi} in that set, find the vertex in V which corresponds to the pair (x_{yi}, y). E'_y is then a set of edges which creates a cycle among the vertices obtained in the above manner from clause y.

Now take

$$E = \left(\bigcup_{x \in X} E_x\right) \cup \left(\bigcup_{y \in Y} E'_y\right)$$

and let $G = (V, E)$ be the graph used for the derived COD instance. As an example, Figure 1 shows such a graph constructed for the 3-SAT formula $(a + b + d')(a + b' + d)(a' + b' + d)$, where numbers are used to indicate the clause associated with each vertex.

To complete the reduction, take $\phi(v) = \theta(e) = 1$, let $q = \frac{1}{2} + \frac{1}{6|V|}$, and let $c' = |X||Y|$. To see this is a valid reduction of 3-SAT to COD, note that the reduction can be easily done in polynomial time, and all that remains is to show that any solution of the COD instance is equivalent to some solution of the 3-SAT instance and vice versa. We use the following lemma:

Lemma 1. *If (ϕ', θ') is a valid solution to the COD instance obtained by the given reduction from 3-SAT, then, for any $x \in X$, exactly one of the following must hold:*

- $\phi'(v_{x,y}) = 2$ *and* $\phi'(v_{x',y}) = 1$ *for all* $y \in Y$
- $\phi'(v_{x,y}) = 1$ *and* $\phi'(v_{x',y}) = 2$ *for all* $y \in Y$

Proof of Lemma 1. Consider the subgraph $G_x = (V_x, E_x)$. Since G_x is a complete bipartite graph, there are $|Y|$ disjoint 2-cycles in G_x, where each such 2-cycle $K_{x,y}$ is comprised of the vertices $\{v_{x,y}, v_{x',y}\}$ and the edges between them. Since each such $K_{x,y}$ has an original throughput $T(K_{x,y}, \phi, \theta) = \frac{1}{2}$, and (ϕ', θ') is a valid solution to COD so that $T(K_{x,y}, \phi', \theta') \geq q > \frac{1}{2}$, then at least one of the following four conditions must hold:

- $\phi'(v_{x,y}) \geq 2$
- $\phi'(v_{x',y}) \geq 2$
- $\theta'\left((v_{x,y}, v_{x',y})\right) = 0$
- $\theta'\left((v_{x',y}, v_{x,y})\right) = 0$

Since there are exactly $|X||Y|$ disjoint 2-cycles $K_{x,y}$ over the entire graph G, then exactly one of the above must hold for each $K_{x,y}$. If this were not true, then

$$c = \sum_{v \in V} \left(\phi'(v) - \phi(v)\right) + \sum_{e \in E} \left(\theta(e) - \theta'(e)\right) > |X||Y|$$

so $c > c'$ and thus (ϕ', θ') cannot be a valid solution, a contradiction.

Furthermore, if $\phi'(v_{x,y}) \geq 2$, then $\phi'(v_{x,y}) = 2$, and if $\phi'(v_{x',y}) \geq 2$, then $\phi'(v_{x',y}) = 2$. If this were not the case, then we obtain $c > |X||Y|$, a contradiction.

For the subgraph G_x let

$$c_x = \sum_{v \in V_x} \left(\phi'(v) - \phi(v)\right) + \sum_{e \in E_x} \left(\theta(e) - \theta'(e)\right)$$

Note $c_x \geq |Y|$, since G_x contains $|Y|$ cycles from among the $K_{x,y}$. But $\sum_{x \in X} c_x \leq c \leq |X||Y|$, so, by Dirichlet's Box Principle, we must have equality with $c_x = |Y|$.

Now if $\phi'(v_{x,y}) = 1$ for some $y \in Y$, then we must have $\phi'(v_{x',y_1}) = 2$ for all $y_1 \in Y$. To see this, assume $\phi'(v_{x,y}) = 1$ and take any $y_1 \in Y$, $y_1 \neq y$. Consider the cycle K' on the vertices $\{v_{x,y}, v_{x',y_1}\}$. Since $y_1 \neq y$, K' shares no edge with any of the $K_{x,y}$, and since $c_x = |Y|$, then $\theta'\left((v_{x,y}, v_{x',y_1})\right) = \theta'\left((v_{x',y_1}, v_{x,y})\right) = 1$. If this were not true, then $c_x > |Y|$, a contradiction. But for cycle K' we also must have $T(K', \phi', \theta') \geq q > \frac{1}{2}$. Thus $\phi'(v_{x',y_1}) = 2$ for all $y_1 \neq y$. Now suppose $\phi'(v_{x',y}) = 1$. By a symmetric argument to the one above, we must have $\phi'(v_{x,y_1}) = 2$ for all $y_1 \neq y$. But now $\phi'(v_{x,y_1}) = \phi'(v_{x',y_1}) = 2$ for all $y_1 \neq y$, so $c_x \geq 2|Y| - 2$, which is a contradiction, since $c_x = |Y|$ and we have $|Y| \geq 3$. Thus, by contradiction, our supposition must have been incorrect, and so $\phi'(v_{x',y}) = 2$.

Thus if $\phi'(v_{x,y}) = 1$ for some $y \in Y$, then $\phi'(v_{x',y_1}) = 2$ for all $y_1 \in Y$. By symmetry, if $\phi'(v_{x',y}) = 1$ for some $y \in Y$, then $\phi'(v_{x,y_1}) = 2$ for all $y_1 \in Y$. Thus either $\phi'(v_{x,y}) = 2$ or $\phi'(v_{x',y}) = 2$ for all $y \in Y$. But, if both these cases hold, then $c_x > |Y|$, a contradiction. Thus exactly one must hold, and Lemma 1 is proven. $\square$

Note that Lemma 1 implies that, given a valid solution (ϕ', θ') to the generated COD instance, we must have $\theta'(e) = 1$ for all $e \in E$. If this were not the case, then $c > |X||Y|$ and we have a contradiction, since $c \leq c'$ for a valid solution. Also, any cycle K must contain some vertex v such that $\phi'(v) = 2$. Otherwise, $T(K, \phi', \theta') = \frac{1}{2}$, a contradiction.

Now consider any valid solution (ϕ', θ') to the generated COD instance. From (ϕ', θ') we derive a variable assignment for the original 3-SAT instance as follows: for every variable $x \in X$, if $\phi'(v_{x,y}) = 2$ for all $y \in Y$, set the value of variable x to 1; otherwise set it to 0. We can see that this solution is indeed a valid solution to the original 3-SAT instance. Let K_y be the 3-cycle generated by the edges of E'_y and the incident vertices. Since (ϕ', θ') is a valid solution to the COD instance, each of the cycles K_y must contain some vertex v such that $\phi'(v) = 2$. There are two cases. Case 1 is where v is of the form $v_{x,y}$ for some $x \in X$. By the construction of K_y, clause y in the 3-SAT formula must contain the literal x. Also, $\phi'(v_{x,y}) = 2$ and thus we have set the value of the variable x to 1. Thus clause y takes the value 1. Case 2 is where v is of the form $v_{x',y}$ for some $x \in X$. By the construction of K_y, clause y in the 3-SAT formula must contain the literal x'. Also, $\phi'(v_{x',y}) = 2$, so $\phi'(v_{x,y}) = 1$ by Lemma 1, and thus we have set the value of the variable x to 0. Thus clause y takes the value 1. Therefore all clauses are satisfied under the variable assignment given above, and hence the above process derives a valid solution for the original 3-SAT instance.

For the converse, consider any variable assignment which satisfies the original 3-SAT instance. From this assignment, a solution to the reduced COD instance can be obtained: if a variable x is assigned the value 1, set $\phi'(v_{x,y}) = 2$ and $\phi'(v_{x',y}) = 1$ for all $y \in Y$; otherwise set $\phi'(v_{x,y}) = 1$ and $\phi'(v_{x',y}) = 2$ for all $y \in Y$. Set $\theta' = 1$. Now we must show this solution is indeed valid. First we show that any cycle K must contain some vertex v such that $\phi'(v) = 2$. If K contains an edge from some E_x, $x \in X$, then K contains both v_{x,y_1} and v_{x',y_2} for some $y_1, y_2 \in Y$. By our construction of ϕ, one of these vertices must satisfy $\phi'(v) = 2$. Otherwise K contains edges only from among $\left(\bigcup_{y \in Y} E'_y\right)$. But since the E'_y induce disjoint 3-cycles, K must contain exactly one such 3-cycle, and hence K contains all the edges from a single E'_y. Since the original 3-SAT variable assignment must satisfy clause y, some literal in clause y must take on the value 1. WLOG let this literal be x. Then, by construction, $\phi'(v_{x,y}) = 2$, and also E'_y contains an edge incident to $v_{x,y}$. Thus K contains a vertex v such that $\phi'(v) = 2$.

Now suppose K contains m vertices for which $\phi'(v) = 2$, and that K contains n vertices in total. Then

$$\begin{aligned}
T(K, \phi', \theta') &= \frac{\sum_{v \in K} \phi'(v)}{\sum_{v \in K} \phi'(v) + \sum_{e \in K} \theta'(e)} \\
&= \frac{n+m}{(n+m)+n} \\
&= \frac{1}{2} + \frac{m}{4n+2m} \\
&\geq \frac{1}{2} + \frac{1}{6n} \\
&\geq \frac{1}{2} + \frac{1}{6|V|}
\end{aligned}$$

But then $T(K, \phi', \theta') \geq q$, and also $c = |X||Y| = c'$, so (ϕ', θ') is a valid solution to the COD instance.

Thus a solution to the original 3-SAT instance exists iff there exists a solution to the reduced COD instance, so the given reduction from 3-SAT to COD is a valid reduction, and hence COD is NP-complete. □

3 Additional Note

Note that the proof for NP-completeness holds even when suitable bounds are placed on the values of $\phi, \theta, \phi', \theta', c'$ and q; hence COD is NP-complete in the strong sense, and there does not exist a pseudo-polynomial time algorithm for COD (unless P=NP).