Operations Research Letters 43 (2015) 586–591

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

Using shortcut edges to maximize the number of triangles in graphs

Sina Dehghani [b], Mohammad Amin Fazli [a,*], Jafar Habibi [a], Sadra Yazdanbod [c]

[a] Computer Engineering Department, Sharif University of Technology, Iran
[b] School of Computer Science, University of Maryland at College Park, United States
[c] School of Computer Science, Georgia Institute of Technology, United States

ARTICLE INFO

Article history:
Received 30 October 2014
Received in revised form
1 September 2015
Accepted 1 September 2015
Available online 16 September 2015

Keywords:
Triangles
Graphs
NP-complete
Approximation algorithm

ABSTRACT

In this paper, we consider the following problem: given an undirected graph $G = (V, E)$ and an integer k, find $I \subseteq V^2$ with $|I| \leq k$ in such a way that $G' = (V, E \cup I)$ has the maximum number of triangles (a cycle of length 3). We first prove that this problem is NP-hard and then give an approximation algorithm for it.

1. Introduction

The problem of augmenting networks in order to optimize their properties has been extensively tackled with a number of different approaches in recent years. The main goal of such augmentation is to improve network efficiency or constructing models with desired properties.

In general, to improve the network efficiency one would have to either change the transmission protocols [17,26,28] or change the underlying structure [4,8,32,9,20]. To support the latter approach, which is the main focus of this paper, an active line of research studies the impact of different structural properties on the performance of different network dynamics [14,29,1,31]. Therefore, to improve different network dynamics, we can optimize their associated structural properties.

The second important application of such network optimization problems, is to calibrate structural network models. These models are simply artificial graphs generated with real network properties and are used as a base for simulating different network dynamics. The main goal of these models is to study network behaviors under different conditions. Although numerous structural network models have been proposed over the years, none of them is complete because each focuses on only a subset of these properties and thus misses the others [30,3,2,18]. When a model N does not satisfy a property P, we can calibrate N by optimizing P with minor modifications to the structure of N.

While heuristics have been applied extensively for a wide range of network properties such as diameter and average path length [23], robustness [16,4,32,21] and synchronizability [8,21], approximation algorithms with guaranteed approximation factors have not received much attention. To the best of our knowledge, the only structural properties for which approximation algorithms and non-approximability results are proposed, are diameter, average path length [20,9,6,7] and Eulerian extension [10,15].

A high density of triangles (a cycle of length 3) is a beneficial structural property of graphs. The main behavior of graphs with this property is their fast collective dynamics [24,27]. Examples of such dynamics can be seen in a wide variety of fields such as relaxation oscillations in gene regulatory networks [19,11], synchronization in biological circuits [14,13], opinion formation in social networks [25] and consensus dynamics of agents in multi agent systems [22]. Thus, optimizing the number of triangles in networks with minor changes in their structure is an important problem.

In this paper, we concentrate on the problem of changing the structure of networks in a way that maximizes the number of their triangles. The change in the structure is done through drawing shortcut edges. We consider the limited budget case where we are only allowed to purchase at most k such shortcut edges.

Definition 1 (Triangle-Max *Problem*)**.** Given an undirected graph $G = (V, E)$ and an integer $k < \binom{|V|}{2} - |E|$. Find a set $I \subseteq V^2$ of at most k shortcut edges ($|I| \leq k$) such that $T(G')$ is maximized,

* Corresponding author.
E-mail address: mohammadamin.fazli@gmail.com (M.A. Fazli).

http://dx.doi.org/10.1016/j.orl.2015.09.003

where $G' = (V, E \cup I)$ and $T(G')$ defines the number of triangles in G'.

In this paper, we first show that Triangle-Max is NP-hard. Then for instances of order n, we give a constant factor approximation algorithm for $k \geq n$ and an $O(n^{\frac{1}{4}})$-factor approximation algorithm for $k < n$.

2. Hardness

In this section, our main goal is to prove the NP-hardness of the Triangle-Max problem. First we define a modified class of the problem in which we only want to maximize the number of triangles with exactly i newly added edges.

Definition 2 (Triangle-Max$^{(i)}$ *(for* $1 \leq i \leq 3$*)*). Let $G = (V, E)$ be an undirected graph and $k < \binom{|V|}{2} - |E|$ be an integer. Find a set $I \subseteq V^2$ of at most k shortcut edges such that $T_i(G', I)$ is maximized where $G' = (V, E \cup I)$ and $T_i(G', I)$ defines the number of i-triangles in G' i.e. the triangles having exactly i edges in I.

Observation 1 shows that the Triangle-Max$^{(1)}$ problem can be solved in polynomial time. A simple greedy algorithm will work for this problem. For each shortcut edge $e = (u, v) \notin E$ with $u, v \in V$, define $F(e)$ to be the set of 1-triangles generated by drawing e. One can see that for each $e \neq e'$, $F(e) \cap F(e') = \emptyset$, therefore selecting k of these shortcut edges with maximum cardinality of F will obtain the optimal solution.

Observation 1. Triangle-Max$^{(1)}$ *is solvable in polynomial time.*

Although Triangle-Max$^{(1)}$ is in P, the other two problems in this class i.e. Triangle-Max$^{(2)}$ and Triangle-Max$^{(3)}$ are both NP-hard. Theorems 1 and 2 prove the hardness of these problems.

Theorem 1. Triangle-Max$^{(2)}$ *is NP-hard.*

Proof. We shall reduce the densest k-subgraph problem (DkS) to Triangle-Max$^{(2)}$ problem. The DkS problem is defined as follows: Given a graph G with n vertices and an integer $k \leq n$, the problem is to find a subgraph of G induced by k of its vertices with maximum number of edges. Let $G = (V, E)$ and k specify an instance of DkS. Assume that $V = \{v_1, v_2, \ldots, v_n\}$ and let x be the number of edges in the densest subgraph of G of size k.

Algorithm 1 Reducing DkS to Triangle-Max$^{(2)}$

---

input: G and k
output: One of the densest subgraphs of G with k vertices
1: Define $V' = \{v'_1, v'_2, \ldots, v'_n\}$ and $U = \{u_1, u_2, \ldots, u_{n^3}\}$. Let $V(G') = V' \cup U \cup \{v\}$.
2: For each $e = v_i v_j \in E(G)$ draw an edge between v'_i and v'_j such that $G'[V']$ and G are isomorphic ($G'[V']$ is the subgraph of G' which is induced by V').
3: Draw an edge between all pairs of vertices in U such that $G[U]$ becomes a clique with n^3 vertices.
4: Insert an edge between every two vertices $v'_i \in V'$ and $u_j \in U$.
5: Set $k' = n^3 + k$.
6: Solve Triangle-Max$^{(2)}$ on the input (G', k'). Let Q be the set of v's neighboring vertices in the returned solution.
7: Define T to be a set of k randomly selected vertices from $Q \setminus U$.
8: **return** vertices in G corresponding to those in T.

Algorithm 1 describes a polynomial-time reduction of the DkS problem to the Triangle-Max$^{(2)}$ problem. In the steps of 1 through 4 of this algorithm, an instance (G', k') of the Triangle-Max$^{(2)}$ problem will be built. G' would be a combination of a clique with

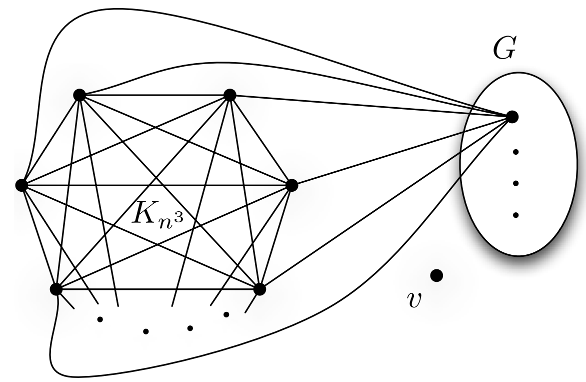

Fig. 1. G' graph.

n^3 vertices, an isomorphic graph to G whose vertices are connected to all vertices of the clique and an isolated vertex v (see Fig. 1).

We claim that any solution of Triangle-Max$^{(2)}$ to the instance (G', k') gives a solution for the DkS problem to the instance (G, k) by the steps 6 through 8 of Algorithm 1.

To prove this, first we need to show that all k' edges in the optimum solution (which we call OPT from now) of Triangle-Max$^{(2)}$ are adjacent to v. Let S be the set of edges in OPT which are not adjacent to v. Edges in S connect vertices in V', so $|S| \leq \binom{n}{2}$. Thus the number of edges adjacent to v is at least $n^3 + k - \binom{n}{2} \geq n^3 - n^2$. There is an optimal solution where all these edges are adjacent to vertices of U. We choose this optimal solution because in this case the maximum number of 2-triangles can be generated. Hence adding an edge to v would increase the number of 2-triangles by at least $n^3 - n^2$.

By removing S's edges, the number of 2-triangles would be decreased by at most $\binom{|S|}{2}$, because each pair of these edges can make at most one 2-triangle. Therefore removing edges in S and adding $|S|$ adjacent edges to v instead, the number of 2-triangles would be increased by at least

$$|S|(n^3 - n^2) - \binom{|S|}{2} \geq |S|\left(n^3 - n^2 - \frac{|S| - 1}{2}\right) \geq n^3 - \frac{3}{2}n^2,$$

which is greater than 0 for $n > 1$. Thus, $S = \emptyset$, i.e. all edges in OPT are adjacent to v.

Now, we prove that the set T returned by Algorithm 1 is one of the densest subgraphs for graph $G'[V']$ (and their corresponding vertices in $V(G)$ for G). Each edge in $G'[Q]$ is included in only one 2-triangle. So the number of 2-triangles created by the edges in OPT is equal to the number of edges in $G'[Q]$. First, notice that there exists a solution for Triangle-Max$^{(2)}$ to the instance (G', k') which creates $y = \binom{n^3}{2} + x + k \cdot n^3$ 2-triangles. It is enough to connect v to the vertices of $U \cup D$ where D is the set of vertices in the densest subgraph of $G'[V']$. We will show that OPT cannot generate more than y 2-triangles and if the equation holds, T must be a densest subgraph of $G'[V']$.

$T \subseteq V'$ and $|T| = k$, therefore the number of edges in $G'[T]$ is less than or equal to x. Also the number of edges in $G'[Q \setminus T]$ is less than or equal to $\binom{n^3}{2}$, because $Q \setminus T$ has exactly $k' - k = n^3$ vertices. Moreover the number of edges between these two subgraphs is less than or equal to $k \cdot n^3$. Thus the number of edges in $G'[Q]$ is less than or equal to y and the equality can only happen when T is a densest subgraph of $G'[V']$, $G'[Q \setminus T]$ is a clique and all vertices in T are connected to all vertices in $Q \setminus T$. □

Our next step is to prove the NP-hardness of the Triangle-Max$^{(3)}$ problem.

Theorem 2. Triangle-Max$^{(3)}$ *is NP-hard.*