Discrete Applied Mathematics ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Discrete Applied Mathematics

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

Graphs that locally maximize clustering coefficient in the space of graphs with a fixed degree sequence☆

Tatsuya Fukami[a], Norikazu Takahashi[b,*]

[a] Department of Informatics, Kyushu University, 744 Motooka, Nishi-ku, Fukuoka, 819-0395, Japan

[b] Graduate School of Natural Science and Technology, Okayama University, 3-1-1 Tsushima-naka, Kita-ku, Okayama, 700-8530, Japan

ARTICLE INFO

Article history:
Received 2 January 2015
Received in revised form 11 April 2015
Accepted 1 October 2016
Available online xxxx

Keywords:
Clustering coefficient
Complex network
2-switch

ABSTRACT

This paper studies the problem of finding graphs that locally maximize the clustering coefficient in the space of graphs with a fixed degree sequence. Such a graph is characterized by the property that the clustering coefficient cannot be increased, no matter how a single 2-switch is applied. First, an explicit formula for the amount of change in the clustering coefficient of a graph caused by a single 2-switch is given. Next, some classes of graphs with the property stated above are presented. An example of such a graph is the one obtained from a tree by replacing its edges with cliques with the same order.

1. Introduction

The clustering coefficient [14] is an important measure that indicates how tightly the neighbors of a vertex in a simple undirected graph are connected to each other. For example, the clustering coefficient for a network of friendship is the probability that two of your friends are themselves friends. It is well known that the clustering coefficient for various networks in the real world tends to be considerably higher than for a random graph with a similar number of vertices and edges [10].

Although its importance is well recognized, the mathematical properties of the clustering coefficient are still not fully understood. In 1999, Watts [13] raised a fundamental question about the clustering coefficient: Among all graphs with given order (i.e., the number of vertices) and size (i.e., the number of edges), which one maximizes the clustering coefficient? As a candidate solution, Watts considered connected caveman graphs and analyzed their clustering coefficient. However, it still remains open whether connected caveman graphs indeed maximize the clustering coefficient [13]. In order to answer this question, Takahashi et al. [11,8,3] have recently addressed the problem of finding graphs with given order and size that maximize the clustering coefficient. Koizuka and Takahashi [8] first considered small graphs and found, by a brute force search, a graph with order n (≤ 10) and size m having the maximum clustering coefficient for each possible pair (n, m). They next used a local search algorithm based on the edge rewiring to find graphs that locally maximize the clustering coefficient. They also proved that the clustering coefficient of any graph composed of two or three cliques sharing one vertex cannot be improved by the algorithm. This result was later generalized by Fukami and Takahashi [3] where it was proved that the clustering coefficient of any graph composed of multiple cliques with orders greater than two sharing one vertex cannot

☆ This work was partially supported by JSPS KAKENHI Grant Number JP15K00035.
* Corresponding author. Fax: +81 86 251 8256.
E-mail address: takahashi@cs.okayama-u.ac.jp (N. Takahashi).

http://dx.doi.org/10.1016/j.dam.2016.10.002
0166-218X/

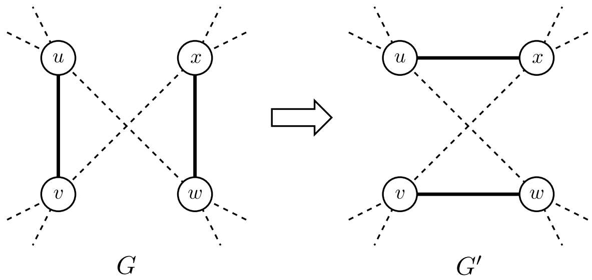

Fig. 1. 2-switch. Dashed lines represent possible edges.

be improved by the algorithm. It was also proved in [3] that the clustering coefficient of any graph obtained from a tree by replacing edges with cliques with the same order other than four cannot be improved by the algorithm.

In this paper, we do not consider the question of Watts but a slightly different one: Among all graphs with given degree sequence, which one maximizes the clustering coefficient? This is closely related to attempts [9,4,6,7,5,12,2] to control the clustering coefficient of scale-free networks [1] without changing the degree distribution. Some of these attempts [9,5,2] are based on 2-switch [15], which is a graph transformation that preserves the degree sequence (see Section 2 for more details). For example, Fukami and Takahashi [2] proposed a local search algorithm based on 2-switch to find a graph such that (i) it has the same degree sequence as a given graph and (ii) the clustering coefficient cannot be increased, no matter how a single 2-switch is applied to it. They also showed experimentally that the clustering coefficient for scale-free networks can be increased to a great extent by their local search algorithm.

We first introduce the notion of the clustering coefficient locally maximizing graph which is defined as a graph such that the clustering coefficient cannot be increased, no matter how a single 2-switch is applied. We then give an explicit formula for the amount of change in the clustering coefficient of a graph caused by a single 2-switch. We finally show some sufficient conditions for a graph to be a clustering coefficient locally maximizing graph. For example, we prove that if a graph is obtained from a tree by replacing its edges with cliques with the same order then it is a clustering coefficient locally maximizing graph.

2. Notations and definitions

Throughout this paper, by a graph, we mean a simple connected undirected graph. A graph is denoted by $G = (V(G), E(G))$ where $V(G)$ is the vertex set and $E(G)$ is the edge set. We assume that vertices of any graph G are labeled from 1 to $|V(G)|$. Then each edge of a graph G is expressed as $\{i, j\}$ where i and j are distinct integers from 1 to $|V(G)|$. For each vertex $i \in V(G)$, the number of vertices $j \in V(G)$ such that $\{i, j\} \in E(G)$ is called the degree of the vertex i and denoted by $d_i(G)$. If G has three vertices i, j and k such that $\{\{i, j\}, \{j, k\}, \{k, i\}\} \subseteq E(G)$ then we say that G has the triangle $\{i, j, k\}$.

Definition 1. The clustering coefficient [14] of a graph G with n vertices is defined as

$$C(G) = \frac{1}{n}\sum_{i=1}^{n} C_i(G).$$

Here $C_i(G)$ is the clustering coefficient of the vertex i which is defined as

$$C_i(G) = \begin{cases} \dfrac{t_i(G)}{d_i(G)(d_i(G)-1)/2}, & \text{if } d_i(G) \geq 2, \\ 0, & \text{otherwise,} \end{cases}$$

where $t_i(G)$ represents the number of unordered pairs of vertices $\{j, k\}$ such that $\{\{i, j\}, \{j, k\}, \{k, i\}\} \subseteq E(G)$, that is, the number of triangles containing the vertex i.

A 2-switch is the replacement of a pair of edges $\{u, v\}$ and $\{w, x\}$ by the edges $\{u, x\}$ and $\{v, w\}$, given that $\{u, x\}$ and $\{v, w\}$ did not appear in the graph originally [15] (see Fig. 1). It is well known that there exists a sequence of 2-switches that can transform a graph G into another graph H if and only if $d_i(G) = d_i(H)$ for all $i \in V(G) = V(H)$ [15]. In the following, when we add an edge $\{i, j\}$ to $G = (V(G), E(G))$ such that $\{i, j\} \notin E(G)$, we express this operation as $E(G) + \{i, j\}$. Similarly, when we remove an edge $\{i, j\}$ from $G = (V(G), E(G))$, we express this operation as $E(G) - \{i, j\}$. Then the 2-switch shown in Fig. 1 is expressed as $E(G') = E(G) - \{u, v\} - \{w, x\} + \{u, x\} + \{v, w\}$.

For any graph G with n vertices, we define the degree sequence as $D(G) = (d_1(G), d_2(G), \ldots, d_n(G))$. Note that entries of $D(G)$ are not sorted in decreasing order. The set of all graphs that have the same degree sequence, say S, is denoted by $\mathcal{G}_S$. Using this notation, we define the clustering coefficient maximizing graph as follows.

Please cite this article in press as: T. Fukami, N. Takahashi, Graphs that locally maximize clustering coefficient in the space of graphs with a fixed degree sequence, Discrete Applied Mathematics (2016), http://dx.doi.org/10.1016/j.dam.2016.10.002