Discrete Applied Mathematics 162 (2014) 202–213

Contents lists available at ScienceDirect

Discrete Applied Mathematics

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

New classes of clustering coefficient locally maximizing graphs☆

Tatsuya Fukami, Norikazu Takahashi *

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

ARTICLE INFO

Article history:
Received 19 September 2012
Received in revised form 12 July 2013
Accepted 24 September 2013
Available online 16 October 2013

Keywords:
Complex network
Clustering coefficient
Connected caveman graph

ABSTRACT

A simple connected undirected graph G is called a clustering coefficient locally maximizing graph if its clustering coefficient is not less than that of any simple connected graph obtained from G by rewiring an edge, that is, removing an edge and adding a new edge. In this paper, we present some new classes of clustering coefficient locally maximizing graphs. We first show that any graph composed of multiple cliques with orders greater than two sharing one vertex is a clustering coefficient locally maximizing graph. We next show that any graph obtained from a tree by replacing edges with cliques with the same order other than four is a clustering coefficient locally maximizing graph. We also extend the latter result to a more general class.

1. Introduction

The clustering coefficient, which was first introduced by Watts and Strogatz [24], is an important measure characterizing large and complex networks in the real world. Roughly speaking, the clustering coefficient is the probability that two vertices adjacent to a given vertex are adjacent to each other. For example, in a network of friendship between individuals, the clustering coefficient represents the probability that two friends of an individual will also be friends of one another [18]. It has been observed that many networks in the real world, such as the Internet, the World Wide Web, networks of coauthorship, metabolic networks and so on, exhibit a high clustering coefficient (see for example [17] and references therein). Also, it has been reported that the clustering coefficient is strongly related to the performance of Hopfield neural networks for associative memories [10], the synchronization of oscillator networks [16], the spread of behavior in online social networks [4] and the evolution of cooperation in games on networks [1,14].

The clustering coefficient is also an important issue in the development of network models. So far, various models have been proposed in order to simulate the behavior of large and complex networks in the real world [2,5,11,18–20,24]. Among them, the preference attachment model proposed by Barabási and Albert [2] is one of the most well-known and widely used models because it exhibits a scale-free degree distribution, which is another important property that can be observed in many networks in the real world. However, it is known that the clustering coefficient of the Barabási and Albert model is very low [3,6]. Therefore, based on this model, many authors have developed scale-free network models with tunable clustering coefficient [8,9,12,21]. In most of these models, the clustering coefficient can be controlled in a certain range by a user-specified parameter. On the other hand, some authors [7,10,15,16] used the 2-switch [25], which rewires two edges simultaneously without changing the degree of each vertex, to increase or decrease the clustering coefficient of a network. In

☆ This work was partially supported by KAKENHI 24560076 and 23310104.

* Correspondence to: Graduate School of Natural Science and Technology, Okayama University, 3-1-1 Tsushima-naka, Kita-ku, Okayama 700-8530, Japan. Tel.: +81 86 251 8179; fax: +81 86 251 8256.

E-mail address: takahashi@cs.okayama-u.ac.jp (N. Takahashi).

0166-218X/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.dam.2013.09.013

particular, Fukami and Takahashi [7] have recently shown experimentally that the clustering coefficient of graphs generated by the Barabási and Albert model can be increased to around 0.8 by applying the 2-switch repeatedly.

As explained above, the importance of the clustering coefficient is widely recognized in the literature. However, properties of the clustering coefficient itself have not been discussed much. To see this, let us consider the following fundamental question: What is the most clustered graph for the given number of vertices and edges? This question was first raised by Watts [22,23]. He considered the connected caveman graph as a candidate solution and derived a general formula for its clustering coefficient. However, it is still not clear whether the connected caveman graph has the highest clustering coefficient or not.

Recently, Koizuka and Takahashi [13] have studied the above-mentioned problem both theoretically and numerically. They first considered small graphs with the number of vertices being less than or equal to 10, and found a graph having the highest clustering coefficient for each possible pair of the number of vertices and the number of edges by using a brute force search. They next applied a local search algorithm to graphs with the number of vertices being less than or equal to 30, and found a graph having a high clustering coefficient for each possible pair of the number of vertices and the number of edges. Their local search algorithm is based on the edge rewiring, that is, the current graph G is replaced with a new graph G' in the neighborhood of G if G' has a higher clustering coefficient than G, where the neighborhood of G is defined as the set of all graphs that can be obtained from G by deleting an edge and adding a new edge. Although this algorithm generates a sequence $G_1, G_2, \ldots$ of graphs such that the sequence $C(G_1), C(G_2), \ldots$ of clustering coefficients is monotone increasing, it is not guaranteed that a graph with the highest clustering coefficient is always reached. In fact, we can easily find a graph G such that its clustering coefficient is higher than any graph in the neighborhood of G but is not the highest among all graphs with the same number of vertices and edges. Koizuka and Takahashi [13] thus focused their attention on such graphs that the clustering coefficient cannot be increased by the local search algorithm, which they call clustering coefficient locally maximizing graphs, and proved that any graph composed of two or three cliques sharing one vertex is a clustering coefficient locally maximizing graph.

The objective of this paper is to find more general classes of clustering coefficient locally maximizing graphs. We first show that any graph composed of multiple cliques with orders greater than two sharing one vertex is a clustering coefficient locally maximizing graph. This is a generalization of the results given by Koizuka and Takahashi [13], but our proof is much simpler than theirs. We next show that any graph obtained from a tree by replacing edges with cliques with the same order other than four is a clustering coefficient locally maximizing graph. We also extend this result to a more general class which includes graphs very similar to connected caveman graphs.

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 a graph $G = (V(G), E(G))$ are always labeled by integers from 1 to $|V(G)|$. Each member of $E(G)$ is thus expressed as $\{i, j\}$ where i and j are distinct integers from 1 to $|V(G)|$. Let $\mathcal{G}(n, m)$ be the set of all graphs composed of n vertices and m edges. Apparently $\mathcal{G}(n, m)$ is non-empty if and only if $n - 1 \leq m \leq n(n - 1)/2$.

The clustering coefficient of a graph can be defined in multiple ways [17,19,24]. In this paper, we focus our attention on the definition introduced by Watts and Strogatz [24]. For a given graph $G = (V(G), E(G)) \in \mathcal{G}(n, m)$, the clustering coefficient of the vertex $i \in V(G)$ is defined by

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

where $d_i(G)$ is the degree of the vertex i and $t_i(G)$ is the number of triangles containing the vertex i, that is,

$$t_i(G) = \left|\{\{j, k\} \in E(G) \mid \{i, j\}, \{i, k\} \in E(G)\}\right|.$$

The clustering coefficient of the graph $G = (V(G), E(G))$ is then defined by

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

If a graph $G \in \mathcal{G}(n, m)$ satisfies $C(G) \geq C(G')$ for all $G' \in \mathcal{G}(n, m)$ then we call G a clustering coefficient maximizing graph in $\mathcal{G}(n, m)$. If a graph $G \in \mathcal{G}(n, m)$ satisfies $C(G) \geq C(G')$ for all $G' \in \mathcal{G}(n, m)$ that are obtained from G by rewiring an edge, that is, removing an edge and adding a new edge, then we call G a clustering coefficient locally maximizing graph in $\mathcal{G}(n, m)$. It is important to note that a clustering coefficient locally maximizing graph in $\mathcal{G}(n, m)$ is not necessarily a clustering coefficient maximizing graph in $\mathcal{G}(n, m)$. In fact, the graph shown in Fig. 1 is a clustering coefficient locally maximizing graph in $\mathcal{G}(6, 7)$ but it is not a clustering coefficient maximizing graph in $\mathcal{G}(6, 7)$ (see [13] for more details).