SindhUniv. Res. Jour. (Sci. Ser.) Vol. 50 (004) 537-540 (2018)

http://doi.org/10.26692/sujo/2018.12.0086

SINDH UNIVERSITYRESEARCH JOURNAL (SCIENCE SERIES)

Computational Complexity of Cycle Discrepancy

L. ASLAM, S. SARWAR, M. M. YOUSAF, S. W. JAFFRY

Punjab University College of Information Technology (PUCIT), University of the Punjab, Lahore.

Received 2nd October 2017 and Revised 12th June 2018

Abstract: The irregularities or the deviations from the state of absolute uniformity are the core focus of the area of discrepancy theory. Cycle discrepancy is about the quantification of the least possible deviation of bi-labeling of graph vertices from an ideal bi-labeling which divides each cycle of a graph in two equal parts. This paper shows that it is NP-hard to compute the cycle discrepancy of a given graph. A polynomial time reduction of Hamiltonian problem to the problem of computing cycle discrepancy is used to establish this result.

Keywords: Hamiltonian; NP-Hard; Cycle Discrepancy

I. INTRODUCTION

Cycle discrepancy is a graph invariant which was first introduced in (Abbasi, and Aslam, 2011). The area of Combinatorics has a branch of discrepancy theory and the concept of cycle discrepancy is inspired from that. The objective of discrepancy theory is to study the deviations from the state of absolute uniformity. Assume we are given a set Q and a set R which contains some subsets of Q. An important question in discrepancy theory is to divide the set Q in a way that every element of R is divided as equally as possible. For the sake of dividing Q into two parts we take two labels say b_1 and b_2 and define a mapping from elements of Q to the set $\{b_1, b_2\}$. Let us call this mapping as a labeling. If we choose a labeling $\beta: Q \rightarrow \{b_1, b_2\}$, then each element of R would also get these labels for their member elements. For each element of R compute the absolute difference between number of elements getting label b_1 and number of elements getting label b_2. The maximum amount of difference obtained would be the discrepancy of the labeling β. In this way every possible labeling would have a value of discrepancy. If we note the minimum value of discrepancy achieved by any possible labeling, it represents the discrepancy of the set system (Q, R). In this setting Q is normally called the ground set. If the set R is obvious from the context, we normally refer the discrepancy of the set system (Q, R) as discrepancy of Q. The reader is referred to (Beck, and Chen, 1987) (Chazelle, 2000) (Matoušek, 1999) (Chen, *et al.*, 2014) and for a further detailed insight about the discipline of discrepancy theory. The foundational work in this area can be found in (Beck and Fiala, 1981) (Spencer, 1985)

Consider a graph $G = (V,E)$, where V denotes the collection of nodes and E represents the collection of edges in the graph. Assume that the ground set is the set of nodes of G that is V and call $C$to be the collection of all cycles in the graph. Take any subset of nodes, say C_i as members of set system (C), if C_i is a cycle in the graph then the discrepancy of the set system (V, C) is known as cycle discrepancy of the graph. Cycle discrepancy of a graph G, is denoted as *cycdisc*(G). It is the nature of set system involved in defining cycle discrepancy due to which graph theoretic tools and methodology is normally useful for investigating this invariant.

A graph is called cubic if every node of the graph has degree exactly three. In (Abbasi, and Aslam, 2011) a tight upper bound of $(n + 2)/6$ is established for cycle discrepancy of a cubic graph with 'n' nodes. It means that a cubic graph on 'n' nodes can have cycle discrepancy of $(n + 2)/6$ at most and further there are such cubic graphs which exhibit this bound.

Cycle discrepancy of three colorable graphs is studied in (Aslam, *et al.*, 2016) and a tight bound of $2 \times \lceil \lfloor n/3 \rfloor / 2 \rceil$ is established. It means, if a graph is three colorable then it cannot have cycle discrepancy more than $2 \times \lceil \lfloor n/3 \rfloor / 2 \rceil$ and there are three colorable graphs which exhibit this bound.

If a cubic graph is also a Toeplitz graph then it is shown in (Aslam, *et al.*, 2018) that it can have cycle discrepancy at most one. Further such cubic Toeplitz graphs are presented which have odd cycles thus the cycle discrepancy of one.

laeeq.aslam@pucit.edu.pk s.sarwar@pucit.edu.pk murtaza@pucit.edu.pk swjaffry@pucit.edu.pk

In (Lovász, 2018) and (Clerk 1973), it is shown that computing combinatorial discrepancy is NP-hard. By using graph theoretic methods, it is shown in this paper that computing cycle discrepancy is also NP-hard. For a given graph to determine that it has a Hamiltonian cycle or not is a famous NP-complete problem. This problem is used in this paper to establish the NP-hardness of the problem of computing cycle discrepancy of a given graph. Reader is referred to (Beck,. and Sós, 1995) and (Stark, 2018) for some latest results showing NP-hardness of problems in different areas.

The next section provides the required definitions, assumptions, and some preliminary work which is useful in establishing the results of this article. The Section 3 provides the argument to establish that computing cycle discrepancy is an NP-hard problem. After that Section 4 briefly concludes the work.

2. DEFINITIONSAND PRELIMINARIES

The concept of cycle discrepancy is defined in for undirected, loop-less graphs without multiple edges. This assumption of simple undirected graphs is also carried in this work. The notation of graph theory used here is inspired from the famous graph theory book (Mesmay, *et al.*, 2018)

A decidable version of a problem is one which can be answered (decided) simply by yes or no. The set of all such problems which can be decided in polynomial time is normally referred as class P. Whereas, set of all such problems whose solution can be verified in polynomial time are normally referred collectively as class NP. If a problem D is member of class NP and all the members of class NP are polynomial time reducible to D then D is called an NP-complete problem. If a problem does not belong to class NP but all members of class NP are polynomial time reducible to it then such problem is referred as an NP-hard problem.

The mapping of the set of nodes of a graph to the set $\{+1, -1\}$ is called labeling of the graph. Instead of using +1 we simply use '+'and instead of using -1 we simply use '$-$', without changing the meaning. Assume a labeling, τ, of a given graph G, then for a subset, M, of nodes of G we can define:

$$\boldsymbol{\tau(M) = \sum_{v \in M} \tau(v)}$$

Note that $\tau^+(M) = \tau(M)$ and $\tau^-(M) = -\tau(M)$.Each cycle in G is defined by a subset of nodes of G. Define a set Y_G containing all cycles of G. Cycle discrepancy of the labeling, τ, is defined as:

$$\boldsymbol{cycdisc(\tau) = \max_{M \in Y_G} |\tau(M)|}$$

For a graph, $G = (V, E)$, the measure of cycle discrepancy which is represented by $cycdisc(G)$ can be defined as:

$$\boldsymbol{cycdisc(G) = \min_{\tau: V \to \{+,-\}} cycdisc(\tau)}$$

It is important to note that cycle discrepancy of any graph which is bipartite will always be zero and the converse is also true. A graph always has cycle discrepancy which is greater or equal to the cycle discrepancy of any of its sub-graphs. That means if a graph has an odd cycle then cycle discrepancy of such graph would be greater than or equal to one.

Theorem 1 from is used in this work. It is reproduced here for the convenience of reader. Here, $\Delta(G)$, denotes maximum degree of a node in the graph, G. The graph constructed in the proof of this theorem is like the one shown in **(Fig. 1)** with t number of triangles. The degree three nodes of triangles in the Fig. 1 are referred as x_i and z_i in the Theorem 1.

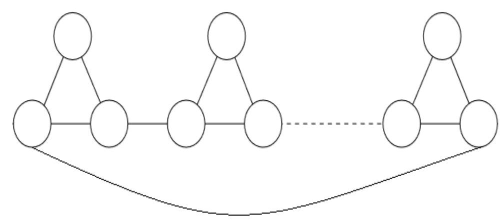

Fig. 1: A chain of t triangles connected in a cycle.

Theorem 1: [Theorem 1]

" For every $n = 3t$ there exist a graph, G, such that $\Delta(G) \leq 3$ and$cycdisc(G) \geq t/2$ "

Proof:

"Let $G_t = (V_t, E_t)$ be a graph on $V_t = \{x_0, y_0, z_0, \cdots, x_{t-1}, y_{t-1}, z_{t-1}\}$ consisting of t triangles connected in a cycle (See Fig. 1). Let $\chi: V_t \mapsto \{+,-\}$ be any labeling of G_t. Consider two cycles C^+ and C^- . The cycle C^+ goes through all the vertices x_i, z_i and also includes all $y_i$such that $\chi(y_i) = +$. Similarly, C^- goes through all the vertices x_i, z_i and includes all the y_i with $\chi(y_i) = -$. Thus $\chi^+(C^+) + \chi^-(C^-) = t$. This implies that

$$\chi^+(\boldsymbol{C^+}) \geq \left\lceil \frac{t}{2} \right\rceil \boldsymbol{or}\ \chi^-(\boldsymbol{C^-}) \geq \left\lceil \frac{t}{2} \right\rceil. \text{"}$$

3. RELATED WORK

To establish the NP-hardness of the problem of computing cycle discrepancy one has to prove that any problem belonging to class NP is reducible to the problem of computing cycle discrepancy in polynomial time. Another way to achieve this task is to show that an established NP-complete problem is polynomial time reducible to the problem of computing cycle discrepancy of a graph. Over here we would use the later approach. The famous NP-complete problem (Garey, and Johnson, 1979), to decide a graph being Hamiltonian or not, would be used. Let R_1 be the Hamiltonian problem which can be formally stated as: given a graph *G*, does *G* has a Hamiltonian cycle? Let R_2 be the decision version of the problem of computing cycle discrepancy which can be stated as: given a graph *G*, does $cycdisc(G) > k$? Where k is a constant. Note that, to establish that R_2 is NP-hard, it is sufficient to show that R_1 is polynomial time reducible to R_2.

Theorem 2:

Given a graph *G*, and a constant *k*, it is NP-hard to decide that $cycdisc(G) > k$.

Proof:

Let $G = (V, E)$be a graph on *n* nodes. We call *G* to be a Hamiltonian graph if it has a Hamiltonian cycle. By using a gadget this graph would be transformed in to $G_T = (V_T, E_T)$ in polynomial time. Then it is shown that *G* is a Hamiltonian graph if

$$\boldsymbol{cycdisc(G_T) \geq \frac{nr}{2} - n.}$$

And in that case *G* is a Hamiltonian graph. This would complete the polynomial time reduction proving that computing cycle discrepancy is NP-hard.

The given graph *G*, has *n* nodes. We take an even number *r* which is very large as compared to *n*. A triangle is a complete graph on three nodes. Make a chain by connecting *r* triangles in a row using extra edges as shown in Fig. 2(a). Now replace every edge of *G* with a chain of *r* triangles to make a new graph named G_T. The graph *G* has *n* nodes while G_T has $n_T = n + 3er$ nodes. The graph *G* has *E* edges while G_T has $E_T = e(3r + 2)$ edges because every edge of *G* is replaced by a chain containing *3r* edges and two edges are used to connect that chain with the original nodes.

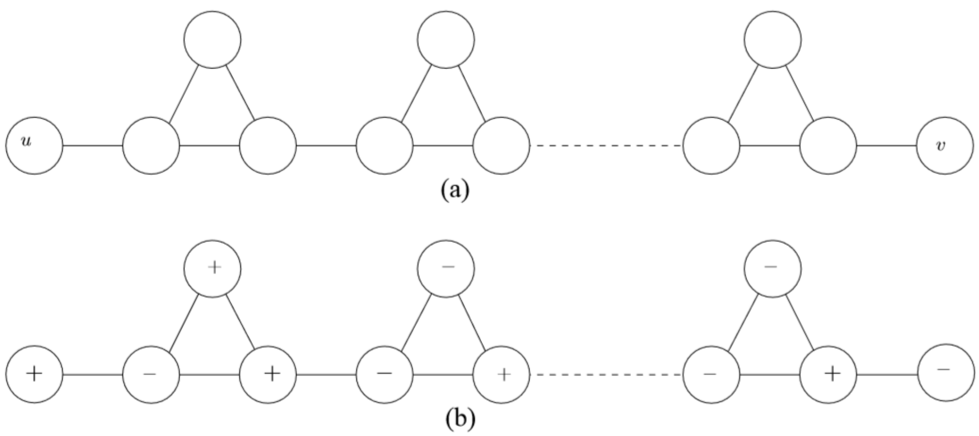

Fig. 2: (a) Chain of *r* triangles used as a gadget to replace edge (*u, v*). (b) The purposed labeling of the chain gadget.

(Fig. 2(a) shows how an edge (*u*, *v*) is substituted by a gadget (chain of triangles).If *G* is a Hamiltonian graph with, C_H, as a Hamiltonian cycle in it then C_H would contain *n* number of edges of *G*. The graph G_T would have a sub-graph corresponding to C_H containing *n* nodes connected by *n* chains each consisting of *r* triangles. Each chain would add at least $r/2$ in the $cycdisc(C_H)$ because of Theorem 1. The cycle discrepancy of C_H would decrease by *n* by non-triangle nodes in the worst scenario. As there are *n* chains in C_H, which is a sub-graph of G_T, we may write:

$$\boldsymbol{cycdisc(G_T) \geq cycdisc(C_H) \geq n\frac{r}{2} - n.}$$

Conversely, if *G* is not a Hamiltonian graph then any cycle in such graph would contain no more than $n - 1$ edges. If we consider the longest possible cycle of *G* then in G_T, the corresponding sub-graph, C_F, would have $n - 1$ nodes connected via $n - 1$ chains each consisting of *r* triangles. Note that in G_T no cycle can contain more than $n - 1$ non-triangle nodes, hence cannot pass through more than $n - 1$ chains. Now we would give a labeling, τ, of G_T. The labeling τ would assign labels to the nodes of G_T as following.

There are n non-triangle nodes. Label '+' to $n/2$ such nodes and label '$-$' to the $n/2$ remaining such nodes. In every chain, nodes having degree three make a path. Label nodes '+' and '$-$' on such path alternatingly. In every chain gadget of G_T, $r/2$ degree two nodes are labeled '+' and the remaining $r/2$ such nodes are labeled '$-$'. An illustration of the labeling τ is shown in Fig. 2(b). It can be easily observed that any cycle in G_T has discrepancy at most

$$\boldsymbol{\frac{(n-1)r}{2} + 1.}$$

For this reason we may write:

$$\boldsymbol{\tau(C_F) < \frac{(n-1)r}{2} + n}.$$

As *r* is very large as compared to *n*, we may write:

$$\boldsymbol{cycdisc(G_T) < \frac{(n-1)r}{2} + n < \frac{nr}{2} - n.}$$

4. CONCLUSIONS

In this paper a polynomial time reduction from Hamiltonian problem to problem of computing cycle discrepancy of a graph is presented. As Hamiltonian problem is a famous NP-complete problem, it is established via this polynomial time reduction that it is NP-hard to compute cycle discrepancy of a given graph.

5. ACKNOWLEDGMENT

The first author is thankful to Dr. Sarmad Abbasi for his time and discussion on this work.All the authors are grateful to Punjab University College of Information Technology for providing support and infrastructure to conduct this research work.

REFERENCES:

Abbasi,A. and L. Aslam, (2011) “The cycle discrepancy of three-regular graphs”, Graphs and Combinatorics, Vol. 27, 27–46,.

Aslam, L., S. Sarwar, M. M. Yousaf, and W. Qounain, (2016) “Cycle discrepancy of d-colorable graphs”, Pakistan Journal of Engineering & Applied Sciences, Vol. 18, 50-55,.

Aslam, L., S. Sarwar, M. M. Yousaf, and S. W. Jaffry, (2018) “Cycle Discrepancy of Cubic Toeplitz Graphs”, Pakistan Journal of Engineering and Applied Sciences, Vol. 22, 14-19,.

Beck, J. and W. L. Chen, (1987) “Irregularities of distribution”, Cambridge University Press, New York.

Bollobás, B. (1988). “Modern Graph Theory” Springer-Verlag,

Beck, J. and W. Fiala, (1981) “Integer making theorems. Discrete Applied Mathematics”, Vol. 3, 1-8.

Chazelle, B. (2000) “The discrepancy method: randomness and complexity”, Cambridge University Press, New York.

Chen, W., A. Srivastav, and G. Travaglini, (2014) “A panorama of discrepancy theory. Lecture notes in mathematics”, Springer International Publishing,.

Clerk M., (1973) A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 68–73,.

Garey, M. R. and D. S. Johnson, (1979) “Computers and Intractability: A Guide to the Theory of NP-Completeness”, A Series of books in the mathematical sciences. W. H. Freeman.

Lovász, L. (2018) “Coverings and colorings of hypergraphs”, Proceedings of the Fourth Southeastern Conference on Combinatorics, Graph Theory, and Computing, Utilitas Math., Winnipeg, Man, 3-12.J.

Matoušek, J. (1999) “Geometric discrepancy: an illustrated guide (Algorithms and Combinatorics)”, Springer, Berlin,.

Mesmay, A., Y. Rieck, E. Sedgwick, and M. Tancer, (2018) “Embeddability in R3 is NP-hard”, Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 1316-1329,.

Spencer, J. (1985) “Six standard deviations suffice”, Transactions of the American Mathematical Society, Vol. 289, .679-706,.

Stark, C. J. (2018) “Learning Optimal Quantum Models is NP-hard”, Physical Review A, Vol. 97, Issue 2, 020103-1-4,