Consecutive ones Block for Symmetric Matrices

Rui Wang, Francis.C.M. Lau*

Department of Computer Science and Information Systems The University of Hong Kong, Hong Kong, P.R. China

Abstract

We show that a cubic graph G with girth $g(G) \geq 5$ has a Hamiltonian Circuit if and only if the matrix $\mathbb{A} + \mathbb{I}$ can be row permuted such that each column has at most 2 blocks of consecutive 1's, where \mathbb{A} is the adjacency matrix of G, \mathbb{I} is the unit matrix, and a block can be consecutive in circular sense, i.e., the first row and the last row are viewed as adjacent rows. Then, based on this necessary and sufficient condition and the NP-completeness of Hamiltonian Circuit for cubic graphs [2], we prove that for every fixed $k \geq 2$ the k-Consecutive Blocks problem (deciding whether a given binary matrix \mathbb{M} can be permuted on rows such that each column has at most k blocks of consecutive 1's) remains NP-Complete even if restricted to (1) symmetric matrices, or (2) matrices having at most 3 blocks of consecutive 1's per row. This result significantly generalizes the related results of [5, 6], and gets its application in [9] in proving the NP-Completeness of all shortest paths Interval Routing Schemes with compactness k for every $k \geq 3$.

Key words: Cubic graph; Hamiltonian graph; Consecutive Blocks; NP-Completeness.

1 The necessary and sufficient conditions

Let G = (V, E) be a simple graph of order n, where $V = \{v_1, v_2, \ldots, v_n\}$, the girth of G is, denoted by g(G), the minimum circuit (cycle) length over all circuits in G; the adjacency matrix of G, denoted by A(G), is a $n \times n$ symmetric binary matrix such that A(i, j) = 1 iif A(i, j) = 1 in column A(i, j) = 1 and either A(i, j) = 1 and A(i, j) = 1 and A(i, j) = 1 and either A(i, j) = 1 and A(i, j) = 1. The Hamiltonicity of a cubic graph A(i, j) = 1 and expressed by A(i, j) = 1 as following.

Theorem 1: A cubic graph G = (V, E) with $g(G) \ge 5$ has a Hamiltonian Circuit if and only if there is a row permutation for $\mathbb{B}(G) = \mathbb{A}(G) + \mathbb{I}$ that results in a matrix having at most 2 blocks of consecutive 1's per column.

^{*}Corresponding author. Email: fcmlau@csis.hku.hk

Proof: Suppose $V = \{v_1, v_2, \dots, v_n\}$, then \mathbb{B} is $n \times n$ symmetric matrix with exactly four 1's per column (row). For any two different rows i and j, and any two different columns k and l, if the 2×2 submatrix they determined has four 1's, i.e., if $\mathbb{B}(i, k) = \mathbb{B}(i, l) = \mathbb{B}(j, k) = \mathbb{B}(j, l) = 1$, then there are three possible cases:

- case 1: none of k and l is i or j, there would be a circuit $v_i v_k v_j v_l v_i$ of length four in G, contradicting $g(G) \geq 5$.
- case 2: one and only one of k and l is i or j, say k = i but $l \neq j$, there would be a triangle $v_i v_j v_k v_i$ in G, again contradicting $g(G) \geq 5$.
- case 3: both k and l are i or j, say k = i and l = j, this is the only possible case for the four entries are all 1's, but must $(v_i, v_j) \in E$ because $\mathbb{B}(i, j) = \mathbb{B}(i, l) = 1$.

From the above analysis, we can infer that making a pair of rows adjacent makes at most two pairs of 1's adjacent in column and it does if and only if the corresponding pair of vertices is adjacent in G. A row permutation defines exactly n pairs of rows adjacent, therefore it makes at most 2n pairs of 1's adjacent in columns.

Noticing that if we want no more than two blocks appear in each column of \mathbb{B} , we need at least 2 pairs of 1's adjacent in each column (because each has four 1's), totally we need to make at least 2n pairs of 1's adjacent in columns. Noticing if a row i is adjacent to two rows j and k such that in G v_i adjacent to v_j and v_k , then column i has at most 2 blocks (3 one entries $\mathbb{B}(j,i)$, $\mathbb{B}(j,j)$, $\mathbb{B}(k,k)$ form one block in column i). So there is a row permutation that achieves at most 2 blocks in each column, if and only if the vertices whose corresponding rows adjacent under the permutation are also adjacent in G, if and only if there is a Hamiltonian Circuit in G. \square

Now, let us turn to blocks consecutive in *linear* sense, in this sense, the number of blocks of consecutive 1's in column j of a binary matrix $\mathbb{A}_{n\times n}$ is defined to be the number of entries $\mathbb{A}(i,j)$ such that $\mathbb{A}(i,j)=1$ and $\mathbb{A}(i+1,j)=0$ or i=n, i.e., we won't take the first row for a natural adjacent row of the last row.

Theorem 2: Let G = (V, E) be a graph with $g(G) \ge 5$, and every vertex $v \in V$ has a degree of 3 except two special vertices $a, b \in V$ which are degree 2, then G has a Hamiltonian Path from a to b if and only if there is a row permutation for $\mathbb{B}(G) = \mathbb{A}(G) + \mathbb{I}$ resulting in a matrix having at most 2 blocks of consecutive 1's per column.

The proof is omitted as it is similar with that of Theorem 1. It is noticeable that both theorems still hold even if the phrase at most is replaced with exactly. Also, the theorem s can be generized to k-regular graph with $g(G) \geq 5$.

2 The NP-Completeness Results

In this section, we will only talk blocks consecutive in linear sense. In next section we will point out that the NP-Completeness results here also hold in the sense of circular blocks.

In 1976, Garey, Johnson and Tarjan [2] have proved that the well-known Hamiltonian Circuit problem remains NP-complete when restricted to graphs which are cubic, triple-connected, planar, and have no face with fewer than 5 edges.

First, we want to make it a little bit stronger by changing the restriction on the size of face to the size of circuit. Starting from an instance graph of this NP-Complete problem, we substitute each vertex with the graph shown in Figure 1 (due to [2]) to get a graph G that obeys $g(G) \geq 5$ while keeping the restrictions. G and the original graph concurrently have or have no Hamiltonian Circuit. Thus, the Hamiltonian Circuit problem remains NP-complete even further restricted to graphs with girth $g(G) \geq 5$.

Then, we would like to reduce it to Hamiltonian Path problem, the problem to decide if there exists a Hamiltonian Path (HP) between two specified vertices. For this, we simply continue to substitute an arbitrarily selected vertex with the graph shown in Figure 2. The resulted graph G has a Hamiltonian Path from a to b iif the original graph has a Hamiltonian Circuit. The graph remains planar, $g(G) \geq 5$, and therefore has no face of size less than 5. but weakens other restrictions a little bit. In conclusion, Hamiltonian Path remains NP-Complete even if restricted to graphs which are

- (1). each vertex is of degree 3 except the two specified vertices a and b which are of degree 2;
- (2). 2-connected; and
- (3). planar; and
- (4). $g(G) \ge 5$ and therefore having no face with fewer than 5 edges.

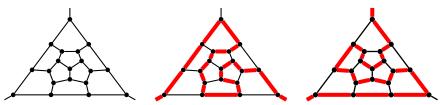


Figure 1. Substitution graph with possible local path (alternate and symmetric paths not shown)

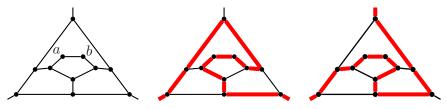


Figure 2. Substitution graph with possible local path (the symmetric path not shown)

By transforming the instance graph G of this restricted Hamiltonian Path to the matrix $\mathbb{B}(G) = \mathbb{A}(G) + \mathbb{I}$, we can prove the following NP-Complete results.

Lemma 1: The 2 Consecutive Block problem for Symmetric matrices (2-CBS for short, defined in the following) is NP-Complete.

Instance: A symmetric binary matrix $\mathbb B$ with at most 3 blocks of consecutive 1's in each row.

Question: Is there a row permutation of $\mathbb B$ resulting in a matrix having at most (or exactly) 2 blocks of consecutive 1's per column?

Proof: From the properties (1) and (4) of G and Theorem 2, it is easy to see the statement holds for the symmetric binary matrix $\mathbb{B}(G)$. So we need only to prove that the matrix $\mathbb{B}(G)$

can be column permuted such that there are at most 3 blocks of consecutive 1's per row, and moreover, it must can be done so in polynomial time.

If the graph G has a matching M (a set of independent edges of G such that each vertex is incident to exactly one edge $e \in M$), the matching M can be found in polynomial time using standard techniques (for example, see [8]), then, for every i the unique j such that $(v_i, v_j) \in M$ can be identified, for each such paired (i, j), making column i and column j adjacent will merge the two one entries $\mathbb{B}(i,i)$ and $\mathbb{B}(i,j)$ into one block in row i and the two one entries $\mathbb{B}(j,i)$ and $\mathbb{B}(j,j)$ into one block in row j, leading to a matrix with at most 3 blocks in each row (because each row of \mathbb{B} has no more than four 1's). Petersen's theorem (Corollary 2.2.3 of [7]) states that every bridgeless cubic graph has a matching (a bridge is an edge $e \in E$ such that G - e is not connected). The graph G here is 2-connected and therefore bridgeless, however, it is not strictly cubic, it is cubic but for two vertices a and b. If we can make Petersen's theorem a bit stronger by claiming that a bridgeless and cubic with at most two exceptions graph G always has a matching, we would have finished. Noticing that the two exception vertices are of degree 2, the proof can be conducted similarly with Petersen's theorem. We present it for self-contained.

We show that such a graph G = (V, E) (bridgeless, cubic with no more than two exceptions) satisfies Tutte's condition $\forall S \subseteq V$, $q(G-S) \leq |S|$, where q(G-S) is the number of odd components of G-S (an odd component is a maximal connected subgraph with odd number of vertices). This is a necessary and sufficient condition for a graph to have a matching [7]. Let $S \subseteq V$ be given, and consider an odd component C of G-S which has no exception vertices. The degrees (in G) of the vertices in C sum to a odd number, but only an even part of this sum arises from the edges of C. So G has an odd number of edges between S and C, and therefore has at least 3 such edges (since G has no bridge). There are at least q(G-S)-2 odd components involving no exception vertices (since there are no more than two exception vertices), and every other component C has at least 2 edges between S to C (again, because G has no bridge). The total number of edges between S and G-S is at least 3(q(G-S)-2)+4>3(q(G-S)-1), implying |S|>q(G-S)-1 (because each vertex is incident with at most 3 edges). \square

We will show for every $k \geq 3$, how to, in polynomial time, increase the number 2 in Lemma 1 to k while keeping the number of blocks no more than 3 per row and the NP-Completeness. This would imply,

Theorem 3 The k Consecutive Block problem for Unsymmetric matrices (k-CBU for short):

Instance: A binary matrix \mathbb{M} with no more than 3 blocks of consecutive 1's per row, and an integer k > 0.

Question: Is there a permutation of rows of \mathbb{M} that results in a matrix having at most k blocks of consecutive 1's per column?

is NP-Complete for every fixed $k \geq 2$.

Further more, before showing the transformation from 2-CBS to k-CBU, we'd like to show a simple transformation from k-CBU to another version which will be defined in Theorem 4.

Starting from matrix \mathbb{M} , an instance of k-CBU, a symmetric matrix \mathbb{M}' can be constructed by joining \mathbb{M} and \mathbb{M}^{T} (the transpose of \mathbb{M}) in the secondary diagonal (shown in the right, where \mathbb{O} is a matrix of appropriate size with all 0 entries). $\mathbb{M}' = \begin{pmatrix} \mathbb{O} & \mathbb{M} \\ \mathbb{M}^{\mathsf{T}} & \mathbb{O} \end{pmatrix}$

Obviously, \mathbb{M}' can be computed from \mathbb{M} in polynomial time, and it can be permuted such that each column has at most k blocks of consecutive 1's if and only if can \mathbb{M} be so. Thus the following theorem is followed.

Theorem 4 The k Consecutive Block problem for Symmetric matrices (k-CBS for short):

Instance: A symmetric binary matrix \mathbb{M} , and an integer k > 0.

Question: Is there a permutation of rows of \mathbb{M} that results in a matrix having at most k blocks of consecutive 1's per column?

is NP-Complete for every fixed $k \geq 2$.

With regard to Theorem 3, when k = 2, it is a special case of Lemma 1, for $k \ge 3$, it can be followed from the following lemma.

Lemma 2 For any $k \geq 3$ we can construct a matrix \mathbb{R} of size $(40(k-2)^3 + 2k - 3) \times 40(k-2)^2$ such that:

- (a). Each row of \mathbb{R} has no more than 2 blocks.
- (b). Under any row permutation, there is a column that has at least k blocks.
- (c). There is a row permutation under which \mathbb{R} has at most k blocks per column.
- (d). There are 2k-3 rows $r_1, r_2, \ldots, r_{2k-3}$ in \mathbb{R} such that any optimal permutation (which achieves no more than k blocks per column) must permute them consecutively in the order (or the reverse order) of $r_1, r_2, \ldots, r_{2k-3}$, that is, permute row r_i adjacent to row r_{i+1} for $1 \leq i \leq 2k-3$.
- (e). Each of the 2k-3 rows $r_1, r_2, \ldots, r_{2k-3}$ has only 1 blocks.

Before proving this lemma, let us show first that Theorem 3 can be proved with it. That is, we can reduce 2-CBS to k-CBU with the help of that particular \mathbb{R} .

Let matrix $\mathbb{B}_{n\times n}$ be any instance of 2-CBS, the instance matrix \mathbb{M} for k-CBU is obtained by joining \mathbb{B} and \mathbb{M} in the main diagonal, and then setting the entries of the k-2 r-rows $r_2, r_4, \ldots, r_{2(k-2)}$ to 1 if those entries at the n columns involving \mathbb{B} . As shown in the following.

$$\mathbb{M} = \begin{pmatrix} \frac{\mathbb{B}}{\mathbf{0}} & \mathbb{O} \\ \mathbf{0} & \mathbf{r}_1 \\ \mathbf{1} & \mathbf{r}_2 \\ \mathbf{0} & \mathbf{r}_3 \\ \vdots & \vdots \\ \mathbf{1} & \mathbf{r}_{2k-4} \\ \mathbf{0} & \mathbf{r}_{2k-3} \\ \hline{\mathbb{O}} & \mathbb{S} \end{pmatrix} \quad \text{where} \quad \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 \\ \vdots \\ \mathbf{r}_{2k-4} \\ \underline{\mathbf{r}_{2k-3}} \\ \hline{\mathbb{S}} \end{pmatrix} = \mathbb{R} \quad \text{and} \quad \begin{cases} \mathbf{1} = (1 \ 1 \ \dots \ 1 \) \\ \mathbf{0} = (0 \ 0 \ \dots \ 0 \) \end{cases}$$

Because of the properties (a) and (e) of \mathbb{R} , and because \mathbb{B} has at most 3 blocks per row, \mathbb{M} has at most 3 blocks per row, so it is an instance of k-CBU. It is easy to see that \mathbb{M} can be constructed in polynomial time. Because of (b), (c) and (d), to permute \mathbb{M} on rows such that

each column has at most k blocks, rows $r_2, r_4, \ldots, r_{2(k-2)}$ must be separated from each other hence form k-2 blocks in each of the n columns of \mathbb{B} , and none of $r_2, r_4, \ldots, r_{2(k-2)}$ is adjacent to any one of the n rows of \mathbb{B} . Thus \mathbb{M} can be row permuted such that each column has at most k blocks if and only if \mathbb{B} can be row permuted such that each column has at most two blocks. So this construction of \mathbb{M} from \mathbb{B} is essentially a polynomial time transformation from 2-CBS to k-CBU, that is, if Lemma 2 is true, then Theorem 3 is true.

Now, what remains for us to do is to prove Lemma 2, which in turn depends on the following lemma.

Lemma 3 For any $k \geq 3$ we can construct a matrix \mathbb{Q} of size $4(k-2)^2 \times 4(k-2)$ such that:

- (i). Each row of \mathbb{Q} has exactly two 1's.
- (ii). Under any row permutation, there is a column that has at least k-1 blocks.
- (iii). There is a row permutation under which \mathbb{Q} has at most k-1 blocks per column.

Proof: Let \mathbb{I} be the $2(k-2) \times 2(k-2)$ unit matrix, then \mathbb{Q} is obtained by vertically juxtaposing 2(k-2) copies of \mathbb{I} and then horizontally appending a matrix obtained by diagonally duplicating $\mathbf{1}^{\mathrm{T}}$ for 2(k-2) times.

$$\mathbb{Q} = \begin{pmatrix} \mathbb{I} & \mathbf{1}^{\mathrm{T}} & \mathbf{0}^{\mathrm{T}} & \dots & \mathbf{0}^{\mathrm{T}} \\ \mathbb{I} & \mathbf{0}^{\mathrm{T}} & \mathbf{1}^{\mathrm{T}} & \dots & \mathbf{0}^{\mathrm{T}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbb{I} & \mathbf{0}^{\mathrm{T}} & \mathbf{0}^{\mathrm{T}} & \dots & \mathbf{1}^{\mathrm{T}} \end{pmatrix} \quad \text{where} \quad \mathbf{1}^{\mathrm{T}} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \quad \text{and} \quad \mathbf{0}^{\mathrm{T}} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Clearly, \mathbb{Q} has $4(k-2)^2$ rows and 4(k-2) columns, with exactly two 1's per row and 2(k-2) 1's per column. (i) is proved.

Any row permutation of \mathbb{Q} makes only $4(k-2)^2-1$ pairs of rows adjacent. Any pair of adjacent rows in \mathbb{Q} has at most one pair of 1's adjacent in column. So, there are totally in \mathbb{Q} at most $4(k-2)^2-1$ pairs of 1's adjacent in column. Among the 4(k-2) columns, there must be a column with less than k-2 pairs of 1's adjacent, implying there are at least k-1 blocks in this column (because there are 2(k-2) 1's in the column). (ii) is derived.

To show (iii), we present a permutation under which each column has no more than k-1 blocks.

Denote the j'th row of the i'th I in Q as $q_{i,j}$ (the 2(k-2)(i-1)+j'th row of Q). The permutation achieving at most k-1 blocks per column arranges the rows in the order of

for
$$j = 1, 2, 3, \dots, 2(k-2) - 1$$
:
 $q_{1,j}q_{1,j+1}q_{2,j+1}q_{2,j}q_{3,j}q_{3,j+1}q_{4,j+1}q_{4,j} \dots q_{2(k-2)-1,j}q_{2(k-2)-1,j+1}q_{2(k-2),j+1}q_{2(k-2),j}$

We leave to the reader the straightforward but tedious verification that under the above order there are k-1 blocks in each of the columns $1, 3, \ldots, 2(k-2)-1$, and k-2 blocks in each of the other columns. The verification will imply (iii). Lemma 3 is followed. \square

Proof of Lemma 2: By diagonally duplicating \mathbb{Q} for 5 times we get \mathbb{Q}_5 , a matrix of size $20(k-2)^2 \times 20(k-2)$ keeping the properties (i)-(iii). Let us first insert two rows r_1, r_2 with all entries equal to 1 into this matrix,

Clearly if we make row r_1 adjacent to row r_2 , we can extend this adjacency to a row permutation under which each column has at most k blocks. On the other hand, if r_1 and r_2

are not adjacent, since each of them can be adjacent to at most two rows in any row permutation, among the five \mathbb{Q} 's, there must be one with no row adjacent to r_1 or r_2 , leading to some column involving this \mathbb{Q} has at least (k-1)+2=k+1 blocks. Thus, row r_1 and row r_2 have to be adjacent if we want no more than k blocks appear in each column. The wanted matrix \mathbb{R} is constructed with the same idea, as shown in the right, which, roughly speaking, is obtained by diagonally duplicating \mathbb{Q}_5 for 2k-4 times, then add 2k-3 rows $r_1, r_2, \ldots, r_{2k-3}$; for every $1 \leq i \leq 2k-4$, for forcing row r_i adjacent to r_{i+1} , set the entries at the columns of the i'th \mathbb{Q}_5 and the rows r_i and r_{i+1} to 1. \mathbb{R}

has $(2k-4)20(k-2)^2 + 2k - 3 = 40(k-2)^3 + 2k - 3$ rows and $(2k-4)20(k-2) = 40(k-2)^2$ columns. It is easy to check M satisfies (a)-(e) to complete the proof of Lemma 2. \Box

3 Discussions and Applications

Lemma 1, Theorem 3 and 4 are stated in the sense of linear blocks, they also hold for circular blocks. For Lemma 1, the circular sense comes from a combination of Theorem 1 and the NP-Completeness of Hamiltonian Circuit for cubic triple-connected graphs with $g(G) \geq 5$. Theorem 3 and 4 hold also for circular sense because appending a row (and a column if to preserve the symmetry) with all zero entries to M will not change its answer in linear sense, but will make the answer in circular sense same as that in linear sense.

2-CBS is important because the matrices are highly restricted and in the question can ask exactly? as well as $at \ most$?. From it we can infer some other NP-Complete variants, such as, deciding whether a matrix can be row permuted such that the total number of blocks in columns is no more than (or exactly) k (this is the result of [3]), or simultaneously permuted on rows and columns such that the total number of blocks in matrix is no more than (or exactly) k. It is conjecturable that deciding whether a given matrix can be row permuted such that each column has $exactly \ k$ blocks is also NP-Complete for every $k \ge 2$.

Consecutive Blocks for binary matrix is a classical problem and has applications in Information Retrieve, DNA computing, and Interval Routing [3, 6, 5]. Many variants in asking about the total number of consecutive 1's blocks in all columns have been proved to be NP-Complete (for example, see [4], A4). Concerning the maximum number k of blocks over all columns, Booth and Lueker [1] gave a linear algorithm for k = 1; Flammini et al. [5] showed its NP-Completeness for general k even if restricted to matrices having at most k blocks per row (they didn't state this restriction in the paper); Goldberg et al. [6] proved the problem is NP-Complete for every $k \geq 2$. The results in this paper significantly extend the related results in [3, 5, 6]. In [9], by taking the advantage of the symmetry of matrices in k-CBS, we prove that for every $k \geq 3$ deciding if a given graph supports an all shortest paths Interval Routing Scheme with compactness k is NP-Complete, and so is for every $k \geq 4$ deciding if a given graph supports an all shortest paths Strict Interval Routing Scheme with compactness k.

References

- [1] K.S.Booth, and S.Lueker, Linear algorithm to recognize interval graphs and test for consecutive ones property, Proc. 7th Ann. ACM Symp. on Theory of computing. New York, pp255-265, 1975.
- [2] M.R.Garey, D.S.Jonson, and R.E.Tarjan. *The planar Hamiltonian circuit is NP-Complete*. SIAM J. comput. 5, pp.704-714, 1976.
- [3] L.T.Kou. Polynomial complete consecutive information retrieval problems. SIAM J. comput. 6, pp.67-75, 1977.
- [4] M.R.Garey, D.S.Jonson. Computers and intractability. A guide to the theory of NP-Completeness. Freeman, San Francisco, CA, 1979.
- [5] M.Flammini, G.Gambosi, and S.Salomone. *Interval routing schemes*, 12th Ann. Symp. on Theoretical Aspects of Computer Science (STACS), Lecture Notes in Computer Science, vol. 900, Springer, Berlin, pp.279-290, March, 1995
- [6] P.W.Goldberg, M.C.Glumbic, H.Kaplan, and R.Shamir. Four strikes against physical mapping of DNA, Journal of Computational Biology, 2(1), pp. 139-152, 1995.
- [7] D.Reinhard, Graph Theory, second edition, Springer-Verlag, New York, 1997.
- [8] C.H.Papadimitriou, K.Steiglitz, Combinatorial optimization, Algorithms and complexity, Dover publications, New York, 1998.
- [9] R.Wang, F.Lau, and Y.Y.Liu, NP-Complete Results for All-Shortest Paths Interval Routing, Unpublished manuscript, 2002.