Strongly Connected Spanning Subgraphs with the Minimum Number of Arcs in Quasi-transitive Digraphs

Jørgen Bang-Jensen * Jing Huang † Anders Yeo ‡

March 9, 1999

Abstract

We consider the problem (MSSS) of finding a strongly connected spanning subgraph with the minimum number of arcs in a strongly connected digraph. This problem is NP-hard for general digraphs since it generalizes the hamiltonian cycle problem. We show that the problem is polynomially solvable for quasi-transitive digraphs. We describe the minimum number of arcs in such a spanning subgraph of a quasi-transitive digraph in terms of the path covering number. Our proofs are based on a number of results (some of which are new and interesting in their own right) on the structure of cycles and paths in quasi-transitive digraphs and in extended semicomplete digraphs. In particular, we give a new characterization of the longest cycle in an extended semicomplete digraph. Finally, we point out that our proofs imply that the MSSS problem is solvable in polynomial time for all digraphs that can be obtained from strong semicomplete digraphs on at least two vertices by replacing each vertex with a digraph whose path covering number can be decided in polynomial time.

Keywords: minimum equivalent digraph, hamiltonian cycle, polynomial algorithm, quasi-transitive digraph, extended semicomplete digraph, path factor, cycle factor, path cover, longest cycle.

1 Introduction

We consider the following problem, which we denote by MSSS (Minimum Spanning Strong Subgraph): given a strongly connected digraph D, find a strongly connected

*Department of Mathematics and Computer Science, Odense University, DK-5230, Denmark (email jbj@imada.sdu.dk). The first author wishes to acknowledge financial support from NSERC (under grant 203191-98).

†Department of Mathematics and Statistics, University of Victoria, Victoria, B.C., V8W 3P4, Canada. (email: jing@math.uvic.ca). The second author wishes to acknowledge financial support from NSERC (under grant 203191-98).

‡Department of Mathematics and Statistics, University of Victoria, Victoria B.C., V8W 3P4 Canada (email: yeo@math.uvic.ca). The third author wishes to acknowledge financial support from the Danish Research Council (under grant 9800435)

spanning subgraph D' of D such that D' has as few arcs as possible. This problem, which generalizes the hamiltonian cycle problem and hence is NP-hard, is of practical interest and has been considered several times in the literature, see e.g. [1, 12, 15, 16, 17, 18]. The MSSS problem is an essential subproblem of the so-called *minimum equivalent digraph problem* (in fact, these two problems can be reduced to each other in polynomial time). Here one is seeking a spanning subgraph with the minimum number of arcs in which the reachability relation is the same as in the original graph (i.e. there is a path from x to y if and only if the original digraph has such a path). Since the MSSS problem is NP-hard, it is natural to study the problem under certain extra assumptions. In order to find classes of digraphs for which we can solve the MSSS problem in polynomial time, we must consider classes of digraphs for which we can solve the hamiltonian cycle problem in polynomial time. This follows from the fact that the hamiltonian cycle problem can be solved if we can solve the MSSS problem.

In [17] the MSSS problem was considered for digraphs whose longest cycle has length r for some r. It was shown that if $r \leq 3$, then the problem is polynomial and that it is NP-hard already when $r = 5$.

In this paper we study the MSSS problem for quasi-transitive digraphs. These digraphs have a nice, recursive structure [8], see Theorem 3.4. Using this structure, Gutin [14] proved that the hamiltonian cycle problem is polynomially time solvable for quasi-transitive digraphs. The approach used to solve the hamiltonian cycle problem in [14] involves solving the problem of finding a minimum path cover of a quasi-transitive digraph.

We give a lower bound for the number of arcs in any minimum spanning strong subgraph of an arbitrary given strong quasi-transitive digraph. This bound can be calculated in polynomial time using Gutin's algorithm for finding a hamiltonian cycle in a quasi-transitive digraph. We prove that this lower bound is also attainable for quasi-transitive digraphs [14]. The proof of this uses a new characterization of a longest cycle in an extended semicomplete digraph.

In the last section we point out the our methods imply that the MSSS problem can be solved efficiently for a much larger superclass of semicomplete digraphs than just quasi-transitive digraphs.

We remark that in [9], the MSSS problem was solved for various generalizations of tournaments. In particular polynomial algorithms were given for the classes of extended semicomplete digraphs and semicomplete bipartite digraphs. Furthermore, it was conjectured in [9] that the MSSS problem is also polynomially solvable for general semicomplete multipartite digraphs.

2 Terminology

We shall always use the number n to denote the number of vertices in the digraph currently under consideration. Digraphs are finite, have no loops or multiple arcs. We use $V(D)$ and $A(D)$ to denote the vertex set and the arc set of a digraph D. We shall use $|D|$ (instead of $|V(D)|$) to denote the number of vertices in D. The arc from a vertex x to a vertex y will be denoted by xy. If xy is an arc, then we say that x *dominates* y and y is *dominated* by x. For disjoint subsets $H, K \subset V(D)$ we use the notation $H \Rightarrow K$ to denote that there are no arcs from K to H.

By a *cycle (path,* respectively) we mean a directed (simple) cycle (path, respectively). If R is a cycle or a path with two vertices u, v such that u can reach v on R, then $R[u, v]$ denotes the subpath of R from u to v. A cycle (path) of a digraph D is *hamiltonian* if it contains all the vertices of D. A digraph is *hamiltonian* if it has a hamiltonian cycle.

An (x, y)-*path* is a path from x to y. A digraph D is *strongly connected* (or just *strong*) if there exists an (x, y)-path and a (y, x)-path for every choice of distinct vertices x, y of D. Let U, W be disjoint subsets of $V(D)$. A (U, W)-path is a path $x_1 x_2 \dots x_k$ such that $x_1 \in U, x_k \in W$ and no other x_i belongs to $U \cup W$.

A digraph T is *semicomplete* if it has no pair of non-adjacent vertices. A *tournament* is a semicomplete digraph with no cycles of length 2. It is well known and easy to prove that every semicomplete digraph has a hamiltonian path and that every strong semicomplete digraph has a hamiltonian cycle. A digraph $D = (V, A)$ is *quasi-transitive* if, for any distinct $x, y, z \in V$, the arcs $xy, yz \in A$ implies that there exists an arc between x and z, i.e., $xz \in A$ or $zx \in A$.

Let $D = (V, A)$ be a digraph. Let $U \subseteq V$ and let $W = (V', A')$ be a subgraph of D. We say that W *covers* U if $U \subseteq V'$.

A collection $\mathcal{F}$ of pairwise vertex disjoint paths and cycles of a digraph D is called a *k-path-cycle factor* of D if $\mathcal{F}$ covers $V(D)$ and has exactly $k \geq 0$ paths. $\mathcal{F}$ is called a *k-path factor* if it contains only paths. We shall call a 0-path-cycle factor a *cycle factor*. A *cycle subgraph* is a collection of vertex disjoint cycles. The *path covering number* of a digraph D, denoted $pc(D)$, is the smallest k for which D has a k-path factor.

Let D be a digraph on p vertices $v_1, ..., v_p$ and let $L_1, ..., L_p$ be a disjoint collection of digraphs. Then $D[L_1, ..., L_p]$ is the new digraph obtained from D by replacing each vertex v_i of D by L_i and adding an arc from every vertex of L_i to every vertex of L_j if and only if $v_i v_j$ is an arc of D $(1 \leq i \neq j \leq p)$. Let D and R be digraphs. Then D is an *extension of* R if there is a decomposition $D = R[I_{a_1}, ..., I_{a_r}]$, $r = |V(R)|$, such that each I_{a_i} induces an independent set in D. An *extended semicomplete digraph* is a digraph which is an extension of a semicomplete digraph. Two vertices x and y in an extended semicomplete digraph $D = R[I_{a_1}, ..., I_{a_r}]$ are said to be *similar* if $x, y \in I_{a_j}$ for some j.

Note that in the rest of the paper, whenever we consider a digraph with a decomposition $D = R[L_1, ..., L_{|R|}]$, we shall think of each L_i both as a subset of $V(D)$ and as a subgraph of D. Furthermore we also think of R as a subgraph of D.

3 Results from other papers

In this section we list a number of results which we will use in the next sections.

Lemma 3.1 *[19] Let $D = (V, A)$ be a digraph which has no cycle factor. Then the vertices of D can be partitioned into disjoint sets Y, Z, R_1, R_2 such that the following holds:*

1. *$D\langle Y \rangle$ has no arcs.*

2. $R_1 \Rightarrow Y \cup R_2$ *and* $Y \Rightarrow R_2$.

3. $|Z| < |Y|$.

Theorem 3.2 *[13] A strong extended semicomplete digraph D is hamiltonian if and only if it has a cycle factor. Furthermore, the length of a longest cycle in D is equal to the maximum number of vertices in a cycle subgraph of D.*

Theorem 3.3 *[13] A longest cycle of an extended semicomplete digraph can be found in time $O(n^{\frac{5}{2}})$.*

Theorem 3.4 *[8] Let D be a quasi-transitive digraph on at least 2 vertices. Then the following holds*

1. *If D is not strong, then D can be decomposed as $D = T[W_1, W_2, \ldots, W_{|T|}]$, where T is a transitive digraph with $|T| \geq 2$ and each W_i is a strong quasi-transitive digraph.*

2. *If D is strong, then D can be decomposed as $D = S[W_1, W_2, \ldots, W_{|S|}]$, where S is semicomplete with $|S| \geq 2$ and each W_i is either a single vertex or a non-strong quasi-transitive digraph. Furthermore, if $s_i s_j s_i$ is a cycle of S, then the corresponding W_i, W_j both have just one vertex.*

The following characterization of hamiltonian quasi-transitive digraphs is given implicitly in [14].

Theorem 3.5 *[14] Let D be a strong quasi-transive digraph with decomposition $D = S[W_1, W_2, \ldots, W_s]$, where $s = |S|$. Let $pc(W_i)$ be the path covering number of the quasi-transitive digraph W_i, $i = 1, 2, \ldots, s$. Let $D_0 = S[H_1, H_2, \ldots, H_s]$ be the extended semicomplete digraph obtained by deleting all arcs inside each W_i (that is $|H_i| = |W_i|$). Then D is hamiltonian if and only if D_0 has a cycle subgraph which covers at least $pc(W_i)$ vertices of H_i, $i = 1, 2, \ldots s$.*

Theorem 3.6 *[14] The path covering number $pc(D)$ of a quasi-transitive digraph D can be calculated and a path cover with $pc(D)$ paths constructed in time $O(n^4)$.*

Theorem 3.7 *[14] There is an $O(n^4)$ algorithm which, given a quasi-transitive digraph D, either returns a hamiltonian cycle in D or a proof that no such cycle exists in D.*

Theorem 3.8 *[8] A quasi-transitive digraph $D = S[W_1, W_2, \ldots, W_{|S|}]$ is hamiltonian if and only if it has a cycle factor $\mathcal{C}$ such that no cycle of $\mathcal{C}$ is a cycle of some $D\langle W_i \rangle$.*

4 Longest cycles in extended semicomplete digraphs

In this section we prove a new characterization of a longest cycle in an extended semicomplete digraph. Besides being a very useful tool in our proof of the main result in the next section, this characterization is also of independent interest. In particular, it implies that, up to switching similar vertices, there is only one longest cycle in an extended semicomplete digraph.

Lemma 4.1 *Let D be an extended semicomplete digraph with an independent set I. If $\mathcal{C}$ is a cycle subgraph covering I, then D contains one cycle C which covers I. Furthermore, given $\mathcal{C}$ and I, we can find one cycle covering I in time $O(n)$.*

Proof: By discarding some cycles if necessary, we may assume that every cycle in $\mathcal{C}$ contains a vertex from I. If $\mathcal{C}$ contains at least two cycles, then let C, C' be distinct cycles from $\mathcal{C}$. Let $x \in V(C), y \in V(C')$ be chosen such that $x, y \in I$. Let x^+, y^+ be the successors of x, y on C, C' respectively. Then xy^+ and yx^+ are arcs of D, since x and y are similar and hence $C[x^+, x]C'[y^+, y]x^+$ is a cycle containing precisely the vertices of $V(C) \cup V(C')$. Now the first claim follows easily by induction on the number of cycles in $\mathcal{C}$. The complexity claim follows from the fact that we can merge the two cycles C, C' in constant time. $\diamond$

Lemma 4.2 *If D is an acyclic extended semicomplete digraph, then $pc(D) = \max\{|I| : I$ is an independent set in $D\}$. Furthermore, starting from D, one can obtain a path cover with $pc(D)$ paths by removing the vertices of a longest path $pc(D)$ times.*

Proof: Let k denote the size of a largest independent set in D. Let $D = S[H_1, H_2, \ldots, H_s]$ be the (unique) decomposition of D such that $H_1, H_2, \ldots, H_s$ are independent sets. Since S is semicomplete, it has a hamiltonian path P and since D is acyclic P is also a longest path in D. Note that since D is acyclic, P contains precisely one vertex from each H_i. Now the claim follows by induction on k. $\diamond$

The following lemma is a special case of a more general result for semicomplete multipartite graphs [13]. Note that it also follows from Theorems 3.2 and 4.4

Lemma 4.3 *Let D be a strong extended semicomplete digraph and let C be a longest cycle in D. Then $D - C$ is acyclic.*

The following characterization of a longest cycle in a strong extended semicomplete digraph is a generalization of Theorem 3.2.

Theorem 4.4 *Let D be a strong extended semicomplete digraph with decomposition $D = S[H_1, H_2, \ldots, H_t]$, $t = |S|$. Let m_i, $i = 1, 2, \ldots, t$, denote the maximum number of vertices from H_i which are contained in a cycle subgraph of D. Then every longest cycle of D contains precisely m_i vertices from each H_i, $i = 1, 2, \ldots, t$.*

Proof: Let C be a longest cycle and suppose without loss of generality that C does not use m_1 vertices from H_1. Let m_1' be the number of vertices from H_1 which are contained in C. First observe that C contains at least one vertex from each H_i.

Indeed, if this is not the case, then choose i so that C has no vertex from H_i. Let x be an arbitrary vertex of H_i. If x has arcs to and from C in D, then it is easy to see that x can be inserted between two vertices of C, contradicting the maximality of C. Suppose without loss of generality that $V(C) \Rightarrow x$. Since D is strong, there is an $(x, V(C))$-path $xq_1q_2 \dots q_t$ in D. Let q_t^- be the predecessor of q_t on C. Then $C[q_t, q_t^-]xq_1q_2 \dots q_t$ is a cycle in D, contradicting the maximality of C. It follows that $1 \leq m_1' < m_1$.

By the definition of m_1 and Lemma 4.1, there is some cycle Q which uses m_1 vertices from H_1. Since all vertices in H_1 have the same adjacencies and $m_1' < m_1$, we can choose Q so that it contains all vertices from H_1 that are on C and at least one extra vertex $x \in H_1 - V(C)$. We will also choose Q so that under the assumption above, $|V(Q) \cap V(C)|$ is maximized.

We claim that for every i such that $H_i \cap V(Q) \not\subset V(C)$ we have $H_i \cap V(C) \subset V(Q)$. If this is not the case, then let u be a vertex of H_i which is on Q but not on C and v a vertex of H_i which is on C but not on Q. Since u and v are similar, we can replace u by v and obtain a new cycle Q' containing m_1 vertices of H_1 which has a larger intersection with C, contradicting the choice of Q above.

Now consider the digraph $D' = D\langle V(C) \cup V(Q) \rangle$. It follows from the fact that C has a vertex from each H_i and that all vertices in H_i are similar that the digraph D' is strong. We claim that D' has a factor. If this is not the case then we can apply Lemma 3.1 to get a partition Y', Z', R_1', R_2' satisfying the conditions of the lemma. It follows from the structure of the arcs determined in Lemma 3.1 that every cycle through a vertex in Y' must use a vertex of Z'. Hence there can be no factor which covers all the vertices in Y'. Since Y' is an independent set in the extended semicomplete digraph D' and hence in D, we have $Y' \subset H_i$ for some i.

For every i such that $H_i \cap V(Q) \not\subset V(C)$ we argued above that all vertices in $H_i \cap V(D')$ are on Q. Hence we cannot have $Y' \subset H_i$ for any of these sets. On the other hand, for every j such that $H_j \cap V(Q) \subset V(C)$, we have all vertices of $H_j \cap V(D')$ on the cycle C. This is a contradiction since C contains a vertex from each H_i.

Thus we have shown that the strong extended semicomplete subgraph D' of D has a cycle factor. By Theorem 3.2, D' has a hamiltonian cycle C'. Now we obtain a contradiction to the assumption C was a longest cycle in D. $\diamond$

5 Smallest spanning strong subgraphs of quasi-transitive digraphs

For an arbitrary quasi-transitive digraph D and a natural number k, we define the quasi-transitive digraph $H_k(D)$ obtained from D as follows: Add two sets of k new vertices $x_1, x_2, \dots, x_k, y_1, y_2, \dots, y_k$. Add all possible arcs from $V(D)$ to x_i along with all possible arcs from y_i to $V(D)$, $i = 1, 2, \dots, k$. Finally, add all arcs of the kind x_iy_j, $i, j = 1, 2, \dots, k$. Note that $H_0(D) = D$.

Definition 5.1 *Let D be a strong quasi-transitive digraph and let $\epsilon(D)$ be the smallest $k \geq 0$ such that $H_k(D)$ is hamiltonian.*

Observe that if $\epsilon(D) \geq 1$, then $\epsilon(D)$ is precisely the path cover number of D. Hence we can calculate $\epsilon(D)$ in time $O(n^4)$ using the algorithms of Theorems 3.6 and 3.7. We show below that $n + \epsilon(D)$ is a lower bound for the number of arcs in every spanning strong subgraph of D.

Lemma 5.2 *For every strongly connected quasi-transitive digraph D every spanning strong subgraph of D has at least $n + \epsilon(D)$ arcs.*

Proof: Let D be a strong quasi-transitive digraph with decomposition $D = S[W_1, W_2, \dots, W_s]$, $s = |S| \geq 2$ (compare with Theorem 3.4). Suppose D has a spanning strong subgraph D' with $n + k$ arcs. We may assume (by deleting some arcs if necessary) that no proper subgraph of D' is spanning and strong. It is easy to prove by induction on k that D' can be decomposed into a cycle $P_0 = C$ and k arc-disjoint paths or cycles $P_1, P_2, \dots, P_k$ with the following properties (where D_i denotes the digraph with vertices $\bigcup_{j=0}^{i} V(P_j)$ and arcs $\bigcup_{j=0}^{i} A(P_j)$ for $i = 0, 1, \dots, t$):

1. For each $i = 1, \dots t$: If P_i is a cycle, then it has precisely one vertex in common with $V(D_{i-1})$. Otherwise the end-vertices of P_i are distinct vertices of $V(D_{i-1})$ and no other vertex of P_i belongs to $V(D_{i-1})$.
2. $\bigcup_{j=0}^{t} A(P_j) = A(D')$.

It is easy to see that this decomposition can be started with P_0 as any cycle in D'. It follows that we may choose $C = P_0$ so that

$$V(C) \not\subset W_i \text{ for } i = 1, 2, \dots, s. \tag{1}$$

Now consider D' as a subgraph of $H_k(D)$. By the minimality assumption on D', each P_i has length at least two. It follows that $H_k(D)$ has a cycle factor consisting of C and k cycles of the form $y_i P'_i x_i y_i$, $i = 1, 2, \dots, k$, where P'_i is the path one obtains from P_i by removing the vertices it has in common with $V(D_{i-1})$ (defined above). By (1) and Theorem 3.8, $H_k(D)$ has a hamiltonian cycle and hence $\epsilon(D) \leq k$. $\diamond$

Below we characterize the optimal solution to the MSSS problem for quasi-transitive digraphs and show that the problem is polynomially solvable.

Theorem 5.3 *The minimum spanning strong subgraph of a quasi-transitive digraph has precisely $n + \epsilon(D)$ arcs. Furthermore, we can find such a subgraph in time $O(n^4)$.*

Proof: Let $D = S[W_1, W_2, \dots, W_s]$, $s = |S| \geq 2$, be a strong quasi-transitive digraph. Using the algorithm of Theorem 3.7 we can check whether D is hamiltonian and find a hamiltonian cycle if one exists. If D is hamiltonian, then any hamiltonian cycle is the optimal spanning strong subgraph. Suppose below that D is not Hamiltonian.

Let $D_0 = S[H_1, H_2, \dots, H_s]$ be the extended semicomplete digraph one obtains by deleting all arcs inside each W_i (that is $|H_i| = |W_i|$ and H_i is obtained from W_i by deleting all arcs). By Theorem 3.5, D_0 has no cycle subgraph which covers at least $pc(W_i)$ vertices of each H_i, $i = 1, 2, \dots, s$.

For each $i = 1, 2, \dots, s$, let m_i denote the maximum number of vertices which can be covered in H_i by any cycle subgraph of D_0. According to Theorem 4.4 every longest cycle C in D_0 contains exactly m_i vertices from H_i, $i = 1, 2, \ldots, s$. By Theorem 3.3 we can find C in time $O(n^{\frac{5}{2}})$. Let

$$k = \max\{pc(W_i) - m_i : i = 1, 2, \ldots, s\}. \tag{2}$$

Define the extended semicomplete subgraph D^* of D as $D^* = S[H_1^*, H_2^*, \ldots, H_s^*]$, where H_i^* is an independent set containing $m_i^* = \max\{pc(W_i), m_i\}$ vertices, $i = 1, 2, \ldots, s$. Since vertices inside an independent set are similar we may think of C as a longest cycle in D^* (i.e. C contains precisely m_i vertices from H_i^*, $i = 1, 2, \ldots, s$). By Lemma 4.3 and Lemma 4.2, $D^* - C$ can be covered by k paths $P_1^*, P_2^*, \ldots, P_k^*$. Since $D^* - C$ is acyclic, we may assume (by Lemma 4.2) that P_1^* starts at a vertex x and ends at a vertex y such that x has in-degree zero and y has out degree zero in $D^* - C$. It follows that there is an arc cx from C to x and an arc yc' from y to C in D^* and hence we can glue P_1^* onto C by adding the arcs cx, yc'. Remove P_1^* and its vertices and consider the remaining paths. It follows by induction on k that adding $P_2^*, P_3^*, \ldots, P_k^*$ one by one, using two new arcs each time, we can obtain a spanning strong subgraph D^{**} of D^* with $|V^*| + k$ arcs.

Now we obtain a spanning strong subgraph of the quasi-transitive digraph D as follows: Since $m_i^* \geq pc(W_i)$ for $i = 1, 2, \ldots, s$, each W_i contains a collection of $t_i = m_i^*$ paths $P_{i1}, P_{i2}, \ldots, P_{it_i}$ such that these paths cover all vertices of W_i. Such a collection of paths can easily be constructed from a given collection of $pc(W_i)$ paths which cover $V(W_i)$. Let $x_{i1}, x_{i2}, \ldots, x_{it_i}$ be the vertex set of H_i^*. Replace x_{ij} in D^{**} by the path P_{ij} for each $i = 1, 2, \ldots, s$, $j = 1, 2, \ldots, t_i$. We obtain a spanning strong subgraph D' of D. The number of arcs in D' is

$$\begin{aligned} A(D') &= \sum_{i=1}^{s}(|W_i| - m_i^*) + (|V^*| + k) \\ &= (n - |V^*|) + (|V^*| + k) \\ &= n + k \end{aligned} \tag{3}$$

It remains to argue that D' is smallest possible. By Lemma 5.2, it suffices to prove that $\epsilon(D) \geq k$.

Suppose $\epsilon(D) = r < k$. By Definition 5.1, the quasi-transitive digraph $H_r(D)$ has a hamiltonian cycle C. It follows from the definition of $H_r(D)$ that we can decompose $H_r(D)$ as $H_r(D) = S'[W_1, W_2, \ldots, W_s, I_r, I_r]$, where I_r is an independent set of r vertices and S' is obtained from S by adding two new vertices x, y such that xy is an arc and x is dominated by all vertices of S and y dominates all vertices of S. Let C' be obtained by contracting each subpath of C which lies entirely inside some W_i. Now delete all remaining arcs inside each W_i. The resulting digraph T is extended semicomplete and has a decomposition $T = S'[I_{a_1}, I_{a_2}, \ldots, I_{a_s}, I_r, I_r]$, where each I_{a_j} denotes an independent set on $a_j \geq 1$ vertices. Since inside every W_i, we only contracted subpaths of C, it follows that $a_i \geq pc(W_i)$ for $i = 1, 2, \ldots, s$. Furthermore, C' is a hamiltonian cycle in T.

Remove the vertices $x_1, x_2, \ldots, x_r, y_1, y_2, \ldots, y_r$ from C'. As the only arcs leaving each x_i go to $\{y_1, y_2, \ldots, y_r\}$, this gives us a collection of r paths $P_1, P_2, \ldots, P_r$ cover all vertices in $T^* = S[I_{a_1}, I_{a_2}, \ldots, I_{a_s}]$. Since all vertices inside the same independent set are similar, we can assume that $P_1, P_2, \ldots, P_r$ are paths in D_0 (D_0 was defined in the beginning of the proof). Let i be chosen such that

$$pc(W_i) - m_i = k. \tag{4}$$

Since $a_i \geq pc(W_i)$ and $r < k$ it follows that some P_j contains two vertices of H_i. Note that if $P_j = z_1 z_2 \ldots z_p$ and $a < b$ are indices so that z_a and z_b are similar, then $z_{a+1} \ldots z_{b-1} z_b z_{a+1}$ is a cycle and $z_a z_{b+1}$ is an arc if $b < p$. Thus we can replace P_j by a cycle and a path $P'_j = P_j[z_1, z_a]P_j[z_{b+1}, z_p]$. Clearly we can continue this way (replacing paths in the current collection by a cycle an a path) until every path in the current collection contains at most one vertex from H_i. This shows that D_0 has a cycle subgraph with covers at least $a_i - r \geq pc(W_i) - r > pc(W_i) - k = m_i$ vertices form H_i. However this contradicts the definition of m_i. This contradiction shows that $\epsilon(D) \geq k$ and the optimality of D' follows from Lemma 5.2.

The proof above can easily be turned into an algorithm which finds a minimum spanning strong subgraph of a given quasi-transitive digraph D. The complexity of the algorithm is dominated by the time it takes to find an optimal path cover in each W_i. By Theorem 3.6 this can be done in $O(n^4)$ time. $\diamond$

6 Remarks and open problems

In order to speed up the algorithm implied by the proof of Theorem 5.3, one would need to find a faster algorithm for finding a hamiltonian cycle in a quasi-transitive digraph. One approach (following Gutin's idea in [14]) would be to find a faster algorithm for the path cover number of quasi-transitive digraphs. This as well as finding a completely different method for solving the hamiltonian cycle problem in quasi-transitive digraphs seems to be challenging open problems.

For another paper which makes good use of the nice recursive structure of quasi-transitive digraphs we refer the reader to [6] in which the problem of finding a heaviest cycle (with respect to weights on the vertices) was solved for quasi-transitive digraphs.

Below we point out that the proofs of our theorems imply a polynomial time algorithm for a much larger class of digraphs than just quasi-transitive digraphs. For every natural number t, let ψ_t be the class of all digraphs for which an optimal path cover can be found in polynomial time $O(n^t)$. For every natural number t, let ϕ_t be the class of all digraphs of the form $D = S[H_1, H_2, \ldots, H_s]$, $s = |S| \geq 2$, where S is a strong semicomplete digraph and $H_i \in \psi_t$, $i = 1, 2, \ldots, s$. By Theorem 3.6 the class ϕ_4 contains all quasi-transitive digraphs.

Using the approach used in this paper it is not difficult to prove the following extension of Theorem 3.5.

Theorem 6.1 *Let t be a natural number and let D be a strong digraph from the class ϕ_t with decomposition $D = S[W_1, W_2, \ldots, W_s]$, where $s = |S|$, $W_i \in \psi_t$, $i = 1, 2, \ldots, s$ and S is a strong semicomplete digraph. Let $pc(W_i)$ be the path cover number of the digraph W_i, $i = 1, 2, \ldots, s$. Let $D_0 = S[H_1, H_2, \ldots, H_s]$ be the extended semicomplete digraph obtained by deleting all arcs inside each W_i (that is $|H_i| = |W_i|$). Then D is* *hamiltonian if and only if D_0 has a cycle subgraph which covers at least $pc(W_i)$ vertices of H_i, $i = 1, 2, \ldots s$.*

Gutin's approach to solving the hamiltonian cycle problem for quasi-transtive digraphs easily extends to a proof of the following.

Theorem 6.2 *For every natural number t, the hamiltonian cycle problem is polynomially solvable for digraphs that belong to ϕ_t.*

Let $D = S[H_1, H_2, \ldots, H_s]$ be a digraph in ϕ_t. To find the minimum strong spanning subgraph in D, let D' be the extended semicomplete digraph obtained from D by deleting all arcs within each H_i for $i = 1, 2, \ldots, s$. By Theorem 3.3, we can find a longest cycle C in D'. Let $m_i = |V(H_i) \cap V(C)|$ for $i = 1, 2, \ldots, s$ and let

$$k = \max\{pc(H_i) - m_i : \ i = 1, 2, \ldots, s\}$$

Using a proof analogous to that of Theorem 5.3, we can show that the minimum strong spanning subgraph of D contains $n + k$ arcs when $k \geq 1$ and is a hamiltonian cycle when $k \leq 0$. Combining this with Theorems 6.1 and 6.2 we get

Theorem 6.3 *For every natural number t, the MSSS problem is polynomially solvable for all digraphs in ϕ_t.* $\diamond$

References

[1] A.V. Aho, M. R. Garey and J. D. Ullman. The transitive reduction of a directed graph, *Siam J. Computing* **1(2)** (1972) 131-137.

[2] R.K. Ahuja, T.L. Magnanti and J.B. Orlin, **Network Flows**, Prentice Hall, New Jersey (1993).

[3] H. Alt, N. Blum, K. Melhorn and M. Paul, Computing of maximum cardinality matching in a bipartite graph in time $O(n^{1.5}\sqrt{m/\log n})$. *Inf. Proc. Letters* **37** (1991) 237-240.

[4] J. Bang-Jensen and G. Gutin, Generalizations of tournaments: A survey, *J. Graph Theory*, **28** (1998) 171-202.

[5] J. Bang-Jensen and G. Gutin, On the complexity of hamiltonian path and cycle problems in certain classes of digraphs, *Discrete Applied Mathematics*, to appear.

[6] J. Bang-Jensen and G. Gutin, Vertex heaviest paths and cycles in quasi-transitive digraphs. *Discrete Math.* **163** (1996) 217–223.

[7] J. Bang-Jensen, G. Gutin, and A. Yeo, A polynomial algorithm for the Hamiltonian cycle problem in semicomplete multipartite digraphs. *J. Graph Theory* **29** (1998) 111-132.

[8] J. Bang-Jensen and J. Huang, Quasi-transitive digraphs. *J. Graph Theory* **20** (1995) 141-161.

[9] J. Bang-Jensen and A. Yeo, Strongly connected spanning subgraphs with the minimum number of arcs in semicomplete multipartite digraphs, submitted.

[10] J.A. Bondy and U.R.S. Murty. **Graph Theory with Applications**, MacMillan Press, 1976.

[11] A. Ghouilà-Houri, Caractérisation des graphes non orientès dont on peut orienter les arrêtes de maniere à obtenir le graphe d'un relation d'ordre, *C.R.Acad. Sci.Paris* **254** (1962) 1370-1371.

[12] P. Gibbons, R. Karp, V. Ramachandran, D. Soroker and R. Tarjan, Transitive compaction in parallel via branchings, *J. Algorithms* **12** 110-125.

[13] G. Gutin, Cycles and paths in complete multipartite digraphs, theorems and algorithms: a survey. *J. Graph Theory* **19** (1995) 481-505.

[14] G. Gutin, Polynomial algorithms for finding hamiltonian paths and cycles in quasi-transitive digraphs. *Australasian J. Combin.* **10** (1994) 231-236.

[15] H.T. Hsu, An algorithm for finding a minimal equivalent graph of a digraph, *J. Assoc. Computing Machinery* **22** (1975) 11-16.

[16] S. Khuller, B. Raghavachari and N. Young, Approximating the minimum equivalent digraph, *Siam J. Computing* **24** (1995) 859-872.

[17] S. Khuller, B. Raghavachari and N. Young, On strongly connected digraphs with bounded cycle length, *Disc. Applied Math.* **69** (1996) 281-289.

[18] K. Simon, Finding a minimal transitive reduction in a strongly connected digraph within linear time. In: *Graph Theoretic concepts in computer science (Kerkrade, 1989).* Springer Verlag Berlin (1990) 244-259.

[19] A. Yeo, How close to regular must a semicomplete multipartite digraph be to secure hamiltonicity? *Graphs and Combinatorics*, to appear.