A Note on the Complexity of Maximizing Temporal Reachability via Edge Temporalisation of Directed Graphs

Alkida Balliu, Filippo Brunelli, Pierluigi Crescenzi, Dennis Olivetti, Laurent Viennot

To cite this version:
Alkida Balliu, Filippo Brunelli, Pierluigi Crescenzi, Dennis Olivetti, Laurent Viennot. A Note on the Complexity of Maximizing Temporal Reachability via Edge Temporalisation of Directed Graphs. 2023. hal-04054527

HAL Id: hal-04054527
https://inria.hal.science/hal-04054527
Preprint submitted on 31 Mar 2023
A Note on the Complexity of Maximizing Temporal Reachability via Edge Temporalisation of Directed Graphs

Alkida Balliu¹, Filippo Brunelli², Pierluigi Crescenzi¹, Dennis Olivetti¹, and Laurent Viennot²

¹GSSI, I-67100 L’Aquila, Italy
²Inria, Irif, Université de Paris, F-75013 Paris, France

March 31, 2023

Abstract

A temporal graph is a graph in which edges are assigned a time label. Two nodes u and v of a temporal graph are connected one to the other if there exists a path from u to v with increasing edge time labels. We consider the problem of assigning time labels to the edges of a digraph in order to maximize the total reachability of the resulting temporal graph (that is, the number of pairs of nodes which are connected one to the other). In particular, we prove that this problem is NP-hard. We then conjecture that the problem is approximable within a constant approximation ratio. This conjecture is a consequence of the following graph theoretic conjecture: any strongly connected directed graph with n nodes admits an out-arborescence and an in-arborescence that are edge-disjoint, have the same root, and each spans Ω(n) nodes.

Keywords: temporal graph; temporal path; time assignment; temporal reachability.

1 Introduction

Temporal graphs have received increasing attention over the last two decades [11, 12, 15, 19] and have been defined in several different ways [1, 10, 13, 2, 6, 5, 14] (see [3] for a classification of temporal graphs). Here, we say that a temporal graph \( G = (V, E) \) is a list \( E \) of temporal edges \((u, v, t)\), where \( u, v \in V \) are two nodes of the graph (called, respectively, tail and head of the temporal edge) and

---

*This work was partly supported by the French National Research Agency (ANR) through the projects Multimod (ANR-17-CE22-0016) and Tempogral (ANR-22-CE48-0001).
Figure 1: A digraph $D$ (left), and an edge temporalisation of $D$ (right). The temporal reachability of the resulting temporal graph is 16, since each node is temporally reachable from any other node.

t is the appearing time of the temporal edge. For each temporal edge $(u, v, t)$, we can traverse the edge starting from $u$ at time $t$ and arrive in $v$ at time $t + 1$, which is the arrival time of the temporal edge.

We study a network optimisation problem related to the notion of reachability in temporal graphs. Given a temporal graph $G$, a (temporal) path from a node $u$ to a node $v$ is a sequence $e_1, e_2, \ldots, e_k$ of temporal edges such that the tail of $e_1$ is $u$, the head of $e_k$ is $v$, and, for any $i$ with $1 < i \leq k$, the tail of $e_i$ is equal to the head of $e_{i-1}$ and the appearing time of $e_i$ is greater than the appearing time of $e_{i-1}$. The temporal reachability of $G$ is the number of pairs of nodes $u$ and $v$ such that $v$ is temporally reachable from $u$, that is, there exists a temporal path from $u$ to $v$. The Maximum Reachability Edge Temporalisation ($mret$) problem consists of, given a directed graph (in short, digraph) $D = (V, E)$, find an edge temporalisation $\tau : E \rightarrow \mathbb{N}$ such that the temporal reachability of the resulting temporal graph is maximized. For example, let us consider the digraph shown in the left part of Figure 1. In the right part of the figure, we show an edge temporalisation of a digraph $D$ with four nodes, such that the temporal reachability of the resulting temporal graph is equal to 16, which is clearly the maximum possible temporal reachability.

The $mret$ problem restricted to undirected graphs has been studied in [9], where it is shown that the problem of deciding whether the resulting temporal graph is temporally connected (that is, for any two nodes $u$ and $v$, $v$ is temporally reachable from $u$) is NP-complete (clearly, this implies that the $mret$ problem restricted to undirected graphs is NP-hard). It is also easy to see that the $mret$ problem restricted to undirected connected graphs can be approximated within a constant approximation ratio, since this simply requires to look for a “centroid” in a spanning tree as a temporalisation where half of the nodes can reach the other half can then easily be computed. Note, however, that temporalising a symmetric digraph is not equivalent to temporalising an undirected graph as different times can be assigned to an edge $(u, v)$ and the symmetric edge $(v, u)$.

In the following, we prove that the $mret$ problem is NP-hard, even when restricted to strongly connected digraphs. We will conclude by conjecturing
that the problem is approximable within a constant approximation ratio, and suggesting a graph theory conjecture which could be used to prove the conjecture concerning the approximability of the MRET problem.

Problems similar to the one considered in this paper have already been analysed [13, 16, 7, 8, 17, 20, 18, 4]. For instance, in [17] the authors propose two cost minimization parameters for temporal network design (that is, the maximum number of appearing times of an edge and the total number of appearing times of all edges), and they study the problem of optimizing these parameters subject to some connectivity constraint.

2 Hardness result

The next result shows that there is no polynomial-time algorithm solving the MRET problem, unless P is equal to NP. In the following, we will refer to edge temporalisations as schedules, that is, as orderings of the edges of the digraph. Indeed, one can easily transform an edge temporalisation $\tau$ into an edge temporalisation $\tau'$, where all time labels are pairwise distinct and where the total reachability according to $\tau$ is preserved. We then note that the total reachability according to $\tau'$ depends only on the ordering of the edges according to their time label. Given a digraph $D = (V, E)$ and a schedule $S$, a node $v$ is said to be $S$-reachable from a node $u$ if it is temporally reachable in the temporal graph $G$ induced by $D$ and the temporalisation $\tau_S$ that assigns appearing time $i$ to the $i$th edge of $S$ for $i \in \left\lceil |E| \right\rceil$. The set of nodes $S$-reachable from a node $u$ is denoted as $R_G(u)$. The $S$-reachability of $D$ is defined as the temporal reachability of the temporal graph induced by $\tau_S$.

**Theorem 1.** The MRET problem is NP-hard, even if the digraph $D$ is strongly connected.

**Proof.** We reduce 3-SAT to MRET as follows. Let us consider a 3-SAT formula $\Phi$, with $n$ variables $x_1, \ldots, x_n$ and $m$ clauses $c_1, \ldots, c_m$. Without loss of generality we will assume that each variable appears positive in at least one clause and negative in at least one clause. We first define the unweighted digraph $D = (V, E)$ as the union of the following gadgets (see Figure 2).

**Variable gadgets** For each variable $x_i$ of $\Phi$, $V$ contains the nodes $t_1^i, t_2^i, f_1^i,$ and $f_2^i$ and $E$ contains the edges $(t_1^i, f_2^i)$, $(f_2^i, f_1^i)$, $(f_1^i, t_2^i)$, and $(t_2^i, t_1^i)$.

**Clause gadgets** For each clause $c_j$, $V$ contains the nodes $c_1^j$ and $c_2^j$. If the literal $x_i$ appears in $c_j$, $E$ contains the edges $(c_1^j, t_1^i)$ and $(t_2^i, c_2^j)$, while if the literal $\neg x_i$ appears in $c_j$, $E$ contains the edges $(c_1^j, f_1^i)$ and $(f_2^i, c_2^j)$. Moreover, for each two clauses $c_h$ and $c_j$ with $h \neq j$, $E$ contains the edge $(c_1^h, c_2^j)$ (see the dashed edges in the figure). Finally, for each clause $c_j$, $V$ also contains the nodes $d_1^i$ and $e_1^i$, for $i \in [K]$ (the value of $K$ will be specified later in the proof), and $E$ contains the edges $(d_1^i, c_1^j)$ and $(c_2^j, e_1^i)$, for $i \in [K]$.
Figure 2: An example of the reduction of 3-sat to mret. The 3-sat formula is \((x_1 \lor x_2 \lor x_3) \land (\neg x_1 \lor x_2 \lor x_3) \land (\neg x_1 \lor \neg x_2 \lor \neg x_3)\).

**Block gadget** \(V\) contains the nodes \(u_1, u_2, u_3,\) and \(u_4,\) and the nodes \(b_i,\) for \(i \in [M]\) (the value of \(M\) will be specified later in the proof). \(E\) contains the edges \(\{(b_i, u_1) : i \in [M]\},\) \((u_1, u_2),\) \(\{(u_2, d^j_l) : j \in [m], l \in [K]\},\)
\(\{(e^j_l, u_3) : j \in [m], l \in [K]\},\) \((u_3, u_4),\) and \(\{(u_4, b_i) : i \in [M]\}.\)

Note that \(D\) is strongly connected. Indeed, let us consider the cycles
\[C_{i,j,l,p} = \langle u_1, u_2, d^j_l, e^j_l, f^1_i, f^2_i, e^2_j, c^2_j, e^l_j, u_3, u_4, b_p, u_1\rangle,\]
where \(j \in [m], i\) is such that \(x_i\) is a literal of the clause \(c_j, l \in [K],\) and \(p \in [M],\)
and the cycles
\[C_{i,j,l,p} = \langle u_1, u_2, d^j_l, e^j_l, f^1_i, f^2_i, e^2_j, c^2_j, e^l_j, u_3, u_4, b_p, u_1\rangle,\]
where \( j \in [m], i \) is such that \( \neg x_i \) is a literal of the clause \( c_j, \ l \in [K], \) and \( p \in [M]. \) The union of these cycles contains each node in \( V, \) and each of these cycles contains node \( u_1. \) This proves the strong connectivity of \( D. \)

In the following, \( B \) denotes the set \( \{b_i : i \in [M]\} \) and \( H \) denotes the set of nodes which do not belong to the block gadget, that is, \( H = V \setminus (\{u_1, u_2, u_3, u_4\} \cup B) \) (note that \( |V| = M + |H| + 4 \) and that \( |H| = 2(K + 1)m + 4n). \)

**Activation of pairs of nodes in a variable gadget** Consider a variable \( x_i \) and the associated variable gadget and a schedule \( S \) of the 4 edges associated to the variable gadget. We say that \( S \) activates the pair \((t^1_i, t^2_i)\) (respectively, \((f^1_i, f^2_i)\)) if \( t^2_i \) (respectively, \( f^2_i \)) is \( S \)-reachable from \( t^1_i \) (respectively, \( f^1_i \)) within the gadget, that is, \((t^1_i, f^2_i), (f^1_i, t^1_i), (f^1_i, t^2_i), (t^1_i, f^2_i)\) are scheduled in that order. Note that no schedule can activate both \((t^1_i, t^2_i)\) and \((f^1_i, f^2_i)\) as the edge \((t^1_i, f^2_i)\) is scheduled either before or after the edge \((f^1_i, t^2_i)\).

**Constructing a schedule from a satisfying assignment.** Suppose that there is an assignment \( \alpha \) that satisfies \( \Phi, \) and let us consider the following schedule \( S. \) First we schedule the edges in \( \{(b_i, u_1) : i \in [M]\} \) (in any arbitrary order), then the edge \((u_1, u_2)\), and then the edges in \( \{(u_2, d^1_i) : j \in [m], i \in [K]\} \) (in any arbitrary order). Then we schedule the edges \( \{(d^1_j, c^1_j) : j \in [m], i \in [K]\} \) (in any arbitrary order), and then, the edges going out from the nodes \( c^1_j, \) for \( j \in [m] \) (in any arbitrary order). Then, for each \( i \in [n], \) if \( \alpha(x_i) = \text{True}, \) we schedule the edges \( (t^1_i, f^2_i), (f^1_i, t^1_i), (f^1_i, t^2_i), \) and \( (t^1_i, f^2_i)\) in this order (thus activating \((t^1_i, t^2_i)\)). Otherwise (that is, \( \alpha(x_i) = \text{False}, \)) we schedule the edges \( (f^1_i, t^2_i), (t^1_i, f^1_i), \) \( (t^1_i, f^2_i), \) and \( (f^1_i, t^1_i)\) in this order (thus activating \((f^1_i, f^2_i)\)).

Then we schedule all the edges entering the nodes \( c^2_j, \) for \( j \in [m] \) (in any arbitrary order), and then all the edges going out from the nodes \( c^2_j, \) for \( j \in [m] \) (in any arbitrary order). Finally, we schedule the edges in \( \{(e^1_j, u_3) : j \in [m], l \in [K]\} \) (in any arbitrary order), then the edge \((u_3, u_4),\) and all the edges in \( \{(u_4, b_i) : i \in [M]\} \) (in any arbitrary order). Let \( G \) be the temporal graph induced by \( D \) and the schedule \( S. \)

First observe that, for any clause \( c_j \) with \( j \in [m], \) there exists a literal that satisfies \( c_j \) according to the assignment \( \alpha. \) Let \( x_i \) (respectively, \( \neg x_i \)) be a literal satisfying \( c_j. \) Since \((t^1_j, t^2_j)\) (respectively, \((f^1_j, f^2_j)\)) is activated, there exists a temporal path from \( c^1_j \) to \( c^2_j \) that goes through variable gadget corresponding to \( x_i. \) This means that \( c^2_j \in R_G(e^1_j) \) and that, for \( l, l' \in [K], \) \( e^l_j \in R_G(e^l_j). \) We now prove a lower bound \( L \) on the \( S \)-reachability by showing a lower bound on the number of nodes temporally reachable from each possible source.

- For any \( v \in V \) and for \( i \in [M], v \in R_G(b_i). \) This adds \( M(M + |H| + 4) \) to \( L. \)

- For \( i \in [4] \) and for \( j \in [M], b_j \in R_G(u_i). \) Moreover, \( \{u_1, u_2, u_3, u_4\} \cup H \subseteq R_G(u_1), \{u_2, u_3, u_4\} \cup H \subseteq R_G(u_2), u_3, u_4 \in R_G(u_3), \) and \( u_4 \in R_G(u_4). \) This adds \( 4M + 2|H| + 10 \) to \( L. \)
• For \( j, h \in [m], i, l \in [K], \) and \( p \in [M], c^2_h, c^1_i \in \mathcal{R}_G(d^3_j) \) (because of the above observation) and \( b_p \in \mathcal{R}_G(d^3_j) \). This adds \( Km(M + Km + m) \) to \( L \).

• For \( j, h \in [m], l \in [K], i \in [M], c^2_h, c^1_i, b_i \in \mathcal{R}_G(c^3_j) \). This adds \( m(M + Km + m) \) to \( L \).

• For \( i \in [n], \) there exists \( j \in [m] \) such that \( c_j \) is satisfied by \( \alpha(x_i) \). Hence, for \( p \in [2], l \in [K], \) and \( h \in [M], e^3_j, b_h \in \mathcal{R}_G(t^3_p) \) and \( e^1_i, b_h \in \mathcal{R}_G(f^3_p) \). This adds \( 4n(M + K) \) to \( L \).

• For \( j \in [m], l \in [K], \) and \( h \in [M], e^3_j, b_h \in \mathcal{R}_G(c^3_j) \). This adds \( m(M + K) \) to \( L \).

• For \( j \in [m], l \in [K], \) and \( h \in [M], b_h \in \mathcal{R}_G(e^3_j) \). This adds \( MKm \) to \( L \).

Thus, the \( S \)-reachability is at least

\[
L = M(M + |H| + 4) + (4M + 2|H| + 10) + Km(M + Km + m) + m(M + Km + m) + 4n(M + K) + m(M + K) + MKm.
\]

**Bounding reachability when \( \Phi \) is not satisfiable.** Let us set \( M \) equal to any value greater than \( (|H| + 5)^2 \). We now prove that, if there exists no truth-assignment satisfying the formula \( \Phi \), then no schedule \( S \) can have \( S \)-reachability greater than or equal to \( L \). First notice that if \( S \) assigns to the edge \((u_3, u_4)\) a starting time smaller than the starting time assigned to \((u_1, u_2)\), then the \( S \)-reachability is less than \( L \). This is because, in this case, for \( i, j \in [M] \) with \( i \neq j \), \( b_j \) is not \( S \)-reachable from \( b_i \). Hence, the \( S \)-reachability is bounded by \( U_1 = M(|H| + 4 + 1) + (|H| + 4)(M + |H| + 4) \). This would happen if, for each node \( v \notin B, \mathcal{R}_G(v) = V \). Since \( L > M^2, U_1 = M(|H| + 4 + 1) + (|H| + 4)(M + |H| + 4) = 2M(|H| + 4) + (|H| + 4)^2 + M < M(|H| + 5)^2 \), and \( M > (|H| + 5)^2 \), it holds that \( L > U_1 \). We can then focus on schedules that assign to the edge \((u_1, u_2)\) a starting time smaller than the starting time assigned to the edge \((u_3, u_4)\). Let \( S \) be such a schedule and let \( G \) be the temporal graph induced by \( D \) and \( S \). We now prove an upper bound \( U_2 \) on the \( S \)-reachability by giving an upper bound on the nodes reachable from each possible source. Observe that, for any two nodes \( u \) and \( v, v \) might belong to \( \mathcal{R}_G(u) \) only if in \( D \) there exists a path from \( u \) to \( v \) that does not include the edge \((u_3, u_4)\) before the edge \((u_1, u_2)\).

• For \( i \in [M], |\mathcal{R}_G(b_i)| \leq |V| \). This adds \( M(M + |H| + 4) \) to \( U_2 \).

• For \( i \in [2], |\mathcal{R}_G(u_4)| \leq |V|, \) while \( |\mathcal{R}_G(u_3)| \leq M + 3 \) and \( |\mathcal{R}_G(u_4)| \leq |V| = M + |H| + 4 \).

• For \( j \in [m] \) and \( i \in [K], \) in the best case \( \mathcal{R}_G(d^3_j) \) contains \( d^3_j, c^1_i \), the 12 nodes corresponding to the three variables appearing in \( c_j \), and the nodes in \( \{e^3_h : h \in [m]\} \cup \{e^1_h : h \in [m], l \in [K]\} \cup \{u_1, u_3, u_4\} \cup B \), yielding \( |\mathcal{R}_G(d^3_j)| \leq M + Km + m + 17 \). However, we can show that there exists an index \( j^* \) such that, for \( l, l' \in [K], c^1_{l'}, \notin \mathcal{R}_G(d^3_{l'}) \), implying that the
In summary, we have that if there exists a schedule $S$ such that the $S$-reachability of $D$ is at least $\mathbb{L}$. This completes the proof of the theorem.
3 Conclusion and open problems

In this paper, we have considered the mret problem, that is, the problem of assigning appearing times to the edges of a digraph in order to maximize the total reachability of the resulting temporal graph. We have proved that this problem is NP-hard, even when the digraph is strongly connected. We conjecture that the mret problem can be approximated within a constant approximation ratio. In particular, we conjecture that any strongly connected digraph admits an edge temporalisation with temporal reachability at least equal to \( c \cdot n^2 \) for some constant \( c > 0 \). One way to prove such a statement would be to prove the following interesting graph theory conjecture.

**Almost Spanning Two Rooted-Arborescences conjecture (astra).** Any strongly connected digraph admits an out-arborescence and an in-arborescence that are edge-disjoint, have the same root, and each spans \( \Omega(n) \) nodes.

Note that it is not difficult to prove that the root of the two arborescences mentioned in the astra conjecture cannot be any node in the graph. For example, let us consider the graph shown in Figure 3. In this case, the node \( x_1 \) cannot be the common root of the two arborescences, since the only in-arborescence and the only out-arborescence with root \( x_1 \) share the edge \((x_2, y_2)\), so that one of the two arborescences cannot include more than one node (of course, this example can be generalized to any even number of nodes).

Note also that the astra conjecture is false if we require that the two arborescences span at least \( \frac{n}{2} \) nodes, for any positive constant \( \epsilon \). For example, consider the digraph \( G = (V, E) \) shown in Figure 4 where, for some integer parameter \( k > 0 \), the set of nodes is \( V = \{x, y, x_1, x_2, y_2, x_3, y_3\} \cup \{z_{i,j} \mid 1 \leq i \leq 3, 1 \leq j \leq k\} \), and the set of edges is \( E = \{(x, y)\} \cup \{(y, x_i), (x_i, y_i), (y_i, x), (y_i, z_{i,1}), (z_{i,k}, x_i) \mid 1 \leq i \leq 3\} \cup \{(z_{i,j}, z_{i,j+1}) \mid 1 \leq i \leq 3, 1 \leq j < k\} \). Observe that the total number of nodes is \( n = 3k + 8 \). Let us first give an upper bound on the minimum between the amount of nodes in the in-arborescence and in the out-arborescence in the case where the root is not \( x \) nor \( y \). For any such node, either the in-arborescence or the out-arborescence can contain at most \( k + 3 \) nodes, since the edge \((x, y)\) can be in one arbores-

![Figure 3](image-url)  

Figure 3: An example of a digraph for which only some nodes can be roots of two arborescences each spanning \( \Omega(n) \) nodes.
Figure 4: An example of a digraph for which there are no two edge-disjoint arborescences with a common root and each spanning more than $n/3 + c$ nodes, for some positive constant $c$.

Consider now the case in which either $x$ or $y$ is the root. Let us suppose that the root is $x$ (the other case can be analysed in a similar way). Since, for each $i = 1, 2, 3$, the edge $(x_i, y_i)$ can be in one arborescence only, then either the in-arborescence or the out-arborescence rooted at $x$ can contain at most $n - 2k = k + 8$ nodes. We thus obtained that, in all cases, either the in-arborescence or the out-arborescence is upper bounded by $k + 8 = n/3 + O(1)$.

References


