Article

# Foremost Walks and Paths in Interval Temporal Graphs ${ }^{\dagger}$ 

Anuj Jain ${ }^{1,2, *, \ddagger(1)}$ and Sartaj Sahni ${ }^{2, \ddagger}$<br>1 Adobe Systems Inc., Lehi, UT 84043, USA<br>2 Department of Computer and Information Sciences and Engineering, University of Florida, Gainesville, FL 32611, USA; sahni@cise.ufl.edu<br>* Correspondence: anujjain@ufl.edu or jainanuj99@gmail.com<br>$\dagger$ This paper is an extended version of our paper published in International Conference on Contemporary Computing (IC3-2022), ACM ICSPS.<br>$\ddagger$ These authors contributed equally to this work.


#### Abstract

The min-wait foremost, min-hop foremost and min-cost foremost paths and walks problems in interval temporal graphs are considered. We prove that finding min-wait foremost and min-cost foremost walks and paths in interval temporal graphs is NP-hard. We develop a polynomial time algorithm for the single-source all-destinations min-hop foremost paths problem and a pseudopolynomial time algorithm for the single-source all-destinations min-wait foremost walks problem in interval temporal graphs. We benchmark our algorithms against algorithms presented by Bentert et al. for contact sequence graphs and show, experimentally, that our algorithms perform up to 207.5 times faster for finding min-hop foremost paths and up to 23.3 times faster for finding min-wait foremost walks.


Keywords: interval temporal graphs; contact sequence temporal graphs; foremost walks; min-hop foremost walks; min-wait foremost walks; min-cost foremost walks; NP-hard

Citation: Jain, A.; Sahni, S. Foremost Walks and Paths in Interval Temporal Graphs. Algorithms 2022, 15, 361.
https://doi.org/10.3390/a15100361
Academic Editor: Roberto Montemanni

Received: 16 August 2022
Accepted: 22 September 2022
Published: 29 September 2022
Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.


Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).

## 1. Introduction

Temporal graphs are graphs in which the edges connecting vertices or the characteristics of these edges may change with time. The applications of temporal graphs include the spread of viral diseases, information dissemination by means of physical/virtual contact between people, understanding behavior in online social networks, modeling data transmission in phone networks, modeling traffic flow in road networks, and studying biological networks at the molecular level [1-7].

Two popular categories of temporal or dynamic graphs are contact-sequence (temporal) graphs and interval-temporal graphs. In a contact-sequence graph, each directed edge $(u, v)$ has the label (departure_time,travel_duration), where departure_time is the time at which one can leave vertex $u$ along edge $(u, v)$ and travel_duration is the time it takes to traverse the edge. Therefore, vertex $v$ is reached at time departure_time + travel_time. Note that a contact-sequence graph may have many edges from vertex $u$ to vertex $v$; each edge has a different departure_time. In an interval-temporal graph, each directed edge $(u, v)$ has a label that is comprised of one or more tuples of the form (start_time,end_time, travel_duration), where start_time $\leq$ end_time defines an interval of times at which one can depart vertex $u$. If one departs $u$ at time $t$, start_time $\leq t \leq$ end_time, one reaches $v$ at time $t+$ travel_duration. The intervals associated with the possibly many tuples that comprise the label of an edge $(u, v)$ must be disjoint. Figures 1 and 2 are examples of contact-sequence and interval-temporal graphs, respectively. A (time-respecting) walk in a temporal graph is a sequence of edges with the property that the end vertex of one edge is the start vertex of the next edge (if any) on the walk, each edge is labeled by the departure time from the start vertex of the edge, and the departure time label on each edge is a valid departure time for that edge and is greater than or equal to the arrival time (if any) at the edge's start vertex (A more formal definition is provided in

Section 2). A (time-respecting) path is a walk in which no vertex is repeated (i.e., there is no cycle). $\langle S, 1, A, 2, B>$ and $\langle S, 0, B, 5, C>$ are example walks in the temporal graphs of Figures 1 and 2, respectively. Both walks are also paths.
$(1,1)$


Figure 1. Contact-sequence temporal graph.


Figure 2. Interval-temporal graph.
An application that can be modeled with a contact-sequence graph is a flight network. Each flight has a departure time at which it leaves the originating airport and a certain travel duration before it can reach the destination airport. On the other hand, if we consider the example of a road network, there is no single instance of time when one needs to depart on a given street. There may be different time windows (or time intervals) during which a street may be open for travel. Further, we may define time windows based on travel duration needed to reach from point $A$ to point $B$ on a given street due to different traffic
conditions during different times of the day (such as office hours). Such networks cannot be modeled using contact-sequence graphs, as a given time window represents infinitely many possible departure times and, hence, infinitely many contact sequence edges. It is easy to see that every contact-sequence graph can be modeled as an interval-temporal graph. Further, when time can be discretized, every interval-temporal graph can be modeled as a contact-sequence graph (with potentially an explosion in the number of edges).

The authors of [8-11] focus on finding optimal paths and walks. Optimization criteria such as foremost (arrive at the destination at the earliest possible time), minhop (use the fewest number of hops when going from the source to the destination vertex), shortest (the time taken to go from the source to the destination is minimized), and so on, are considered. Gheibi et al. [12] present a new data structure for contact-sequence graphs that results in faster algorithms for many of the path problems studied in [8]. While [9] focuses on intervaltemporal graphs, refs. [8,10] use the contact-sequence model. Ref. [10] presents an algorithm for finding walks that optimize any linear combination of eight different optimization criteria (foremost, reverse_foremost, fastest, shortest, cheapest, most_likely, min_hop, min_wait). In our earlier paper [13], we developed algorithms to find optimal foremost and min-hop paths in interval-temporal graphs. These algorithms were demonstrated experimentally to run faster than earlier algorithms for these problems.

A temporal graph may have many walks/paths that optimize a criterion such as foremost. In some applications, it is required to find a walk/path that optimizes a secondary criterion from among all walks/paths that optimize a primary criterion. For example, we may desire a min-wait foremost path (i.e., a foremost path in which the sum of the wait times at intermediate vertices is minimized) or a min-hop foremost path (a foremost path that goes through the fewest number of edges). For example, when selecting flights to go from $A$ to $B$ one may wish to use a min-wait foremost path (a route that minimizes the total wait time at intermediate airports while getting to the destination at the earliest possible time) or a min-hop foremost path (a route that involves the fewest number of connections while guaranteeing the earliest possible arrival time). In this paper, we examine the min-wait, min-hop, and min-cost (each edge or edge interval has an additional attribute, its cost), foremost walks/paths problems.

As discussed in Section 6, the algorithm by Bentert et al. can be tuned by carefully choosing the coefficients for different optimization criteria to find min-hop foremost paths and min-wait foremost walks in contact-sequence temporal graphs. We use this method to benchmark our algorithms against the algorithm by Bentert et al. and show that we perform about 207.5 times faster for finding min-hop foremost paths and up to 23.3 times faster for finding min-wait foremost walks. Further, we solve these problems on the intervaltemporal graphs that can represent a wider problem space for temporal graphs, as opposed to the contact-sequence temporal graphs.

Our main contributions in this paper are:

1. We show that the problems for finding min-cost foremost and min-wait foremost paths and walks in interval temporal graphs are NP-hard.
2. We develop a polynomial time algorithm for the single-source all-destinations minhop foremost paths problem in interval-temporal graphs.
3. We develop a pseudopolynomial time algorithm for the single-source all-destinations min-wait foremost walks problem in interval-temporal graphs.
4. We show that the problem of finding min-hop foremost paths and min-wait foremost walks can be modeled using the linear combination formulation employed in Section 6 when the optimization criteria are discrete (e.g., time is discrete). Our modeling methodology readily extends to other primary and secondary criteria as well as to multiple levels of secondary criteria (e.g., shortest min-hop foremost path).
5. We benchmark our algorithms against the algorithm of Bentert et al. [10] using datasets for which the preceding modeling can be used. On these datasets, our algorithm is up to 207.5 times faster for finding min-hop foremost paths and up to 23.3 faster for finding min-wait foremost walks.

The roadmap of this paper is as follows. In Section 2, we describe the problems of finding min-hop foremost paths ( $m h f$ paths), min-wait foremost walks ( $m w f$ walks), and min-cost foremost paths ( $m c f$ paths). In Section 3, we show that the problems of finding $m w f$ paths and walks and the problem for finding $m c f$ paths in interval temporal graphs are NP-hard. In Section 4, we present theorems that describe the properties of min-hop foremost paths in temporal graphs. We review the data structures used to represent interval-temporal graphs along with some fundamental functions necessary for the algorithm for finding $m h f$ paths. Finally, we present the algorithm for finding mhf paths in interval-temporal graphs along with the proof of its correctness and its computational complexity. In Section 5, we present theorems that describe the properties of $m w f$ walks in interval-temporal graphs. We introduce additional data structures required by the algorithm we propose for finding $m w f$ walks in interval-temporal graphs. Finally, we present the algorithm for finding the $m w f$ walks in interval-temporal graphs along with proof of its correctness and complexity analysis. In Section 6, we show how the problems for finding foremost paths and walks with a secondary optimization criteria can be modeled as optimal walks with linear combination of multiple optimization criteria. We use this modeling to find $m h f$ paths and $m w f$ walks in contact sequence graphs using the algorithm of Bentert et al. [10], which optimizes a linear combination of criteria. In Section 7, we compare our algorithms for finding mhf paths and $m w f$ walks with the algorithm by Bentert et al. [10] by transforming contact-sequence graphs to interval-temporal graphs and vice versa. Finally, we conclude in Section 8.

## 2. Foremost Walks and Paths

A walk (equivalently, valid walk, temporal walk or time-respecting walk) in a temporal graph is an alternating sequence of vertices and departure times $u_{t 0}, t_{0}, u_{t 1}, t_{1}, \ldots, u_{t k}$ where (a) $t_{i}$ is a permissible departure time from $u_{t i}$ to $u_{t_{i+1}}$ and (b) for $\left.0 \leq i<k-1, t_{i}+\lambda_{i}\right) \leq t_{i+1}$. Here, $\lambda_{i}$ is the travel duration when departing $u_{t i}$ at $t_{i}$ using the edge $\left(u_{t i}, u_{t_{i+1}}\right)$ (i.e., $t_{i}+\lambda_{i}$ is the arrival time at $u_{t_{i+1}}$ ). For this walk, $u_{t 0}$ is the source vertex and $u_{t k}$ the destination. Note that every walk in which a vertex is repeated contains one or more cycles. A walk that has no cycle (equivalently, no vertex repeats) is a path.
$W_{1}=S, 0, B, 5, C$ is a walk from $S$ to $C$ in the temporal graph of Figure 2. $W_{2}=S, 0, A, 1, B, 3, C, 5, A$ is another walk in the temporal graph of Figure 2. $W_{1}$ does not contain any cycles as none of the vertices are repeated from source to destination. Hence, $W_{1}$ is also a path. However, $W_{2}$ has vertex $A$ that repeats. Therefore, it is not a path. We are interested in foremost paths and walks in a temporal graph that start at a vertex $s$ at a time $\geq t_{\text {start }}$ and end at another vertex $v$. As defined in [8-10], a foremost walk is a walk from a source vertex $s$ to a destination vertex $v$ that starts at or after a given time $t_{s t a r t}$ and has the arrival time $t_{f}$, which is the earliest possible arrival time at the destination $v$ among all possible walks from $s$ to $v$.

### 2.1. Min-Hop Foremost

A min-Hop foremost (mhf) walk is a foremost walk from a source vertex $s$ to another vertex $v$ that goes through the fewest number of intermediate vertices. We observe that every $m h f$ walk is an $m h f$ path, as cycles may be removed from any $s$ to $v$ walk to obtain a path with a fewer number of hops and the same arrival time at $v$. For this reason, we refer to $m h f$ walks as $m h f$ paths in the rest of this paper.

As an example, consider the interval-temporal graph of Figure 3. Every walk in this graph is also a path. The paths $P 1=<a, 0, b, 1, c, 7, d>$ and $P 2=<a, 0, c, 7, d>$ arrive at $d$ at time 8 and are foremost paths from $a$ to $d$. $P 2$ is a 2-hop foremost path while $P 1$ is a 3-hop foremost path. $P 2$ is the only min-hop foremost path or mhf path from $a$ to $d$ in Figure 3.


Figure 3. Example of $m h f$ paths from source vertex $a$.

### 2.2. Min-Wait Foremost

A min-wait foremost or mwf walk is a foremost walk from a source vertex $s$ to any other vertex $v$ that accumulates minimum total wait time at the vertices visited by the walk. The wait time at each vertex $u$ is departure_time ( $u$ ) - arrival_time ( $u$ ). The total wait time accumulated by the walk is the sum of the wait times at each vertex $u$. Therefore, $m w f$ walk is a foremost walk from $s$ to $v$ that minimizes $\left(\sum_{u_{t i}}\right.$ departure_time $\left(u_{t i}\right)$-arrival_time $\left.\left(u_{t i}\right)\right)$ where $u_{t i} \in \operatorname{vertices}(\operatorname{mwf}(s, v))$ and $u_{t i} \neq s ; u_{t i} \neq v$, as there is no wait time accumulated at the source vertex and the destination vertex.
$m w f$ walks can have cycles, as is evident from the example of Figure 4. The $m w f$ walk from source vertex $s$ to destination vertex $b$ is $m w f(s, b)=s, 0, a, 2, c, 4, d, 5, a, 7, b$, arriving at $b$ at time 8 with total wait time of 1 accumulated at vertex $a$. Alternate foremost paths from $s$ to $b$ are $P 1=s, 0, a, 4, b$, arriving at time 8 with a wait time of 3 , and $P 2=s, 0, a, 7, b$, also arriving at 8 with a wait time of 6 . Therefore, to reduce the total wait time along the walk, we may need to go through cycles instead of waiting for a long time at a given vertex.


Figure 4. Example of $m w f$ walks from source vertex $s$.

### 2.3. Min-Cost Foremost

For this problem, we assume there is a non-negative cost associated with every edge traveled along the walk from source vertex $s$ to a destination vertex $v$. The cost of an edge may depend on the departure time from the edge's start vertex. The min-cost foremost or $m c f$ walk is a foremost walk from source vertex $s$ to any other vertex $v$ that incurs minimum cost along the walk. The cost incurred on a time arc from a vertex $u_{t i}$ to $u_{t i+1}$ departing at time $t_{i}$ is denoted by a function $c\left(u_{t i}, u_{t i+1}\right)$. Therefore, the objective for the min-cost foremost or mcf walk problem is to find a walk from a given source vertex $s$ to a destination vertex $v$ with the arrival time $t_{f}$, such that among all the walks from $s$ to $v$ starting at or after $t_{s t a r t}$ and arriving at $t_{f}$, we choose a walk that accumulates minimum cost along the way or minimizes $\left(\sum_{u_{t i}} c\left(u_{t i}, u_{t i+1}\right)\right.$. If there are cycles in a walk from $s$ to $v$, we could eliminate those cycles and arrive at the destination at the same time, or sooner, at a cost that is the same, or less. Therefore, every $m c f$ path is also an $m c f$ walk. For this reason, we refer to $m c f$ walks as $m c f$ paths in the rest of this paper.

## 3. NP-Hard Foremost Path and Walk Problems in Interval Temporal Graphs

Several problems are known to be NP-hard for contact-sequence temporal graphs. For example, Bhadra et al. [14] show that computing several types of strongly connected components is NP-hard; Casteigts et al. [15] show that determining the existence of a nowait path (In a no-wait path, the arrival and departure times at each intermediate vertex are the same.) between two vertices is NP-hard; and Zschoche et al. [16] show that computing several types of separators is NP-hard. Additional complexity results for contact-sequence temporal graphs appear in [15]. Since contact-sequence temporal graphs are a special case of interval-temporal graph, as discussed in Section 1, every problem that is NP-hard for the contact-sequence model remains NP-hard in the interval model. However, the reverse may not be true, as the transformation from the interval model to the contact-sequence model entails a possible explosion in the instance size. In this section, we demonstrate that $m w f$ and $m c f$ path and walk problems are NP-hard in the interval model but polynomially solvable in the contact-sequence model. In fact, in [13], we show that finding a no-wait path from a given source vertex $u$ to a destination vertex $v$ in an interval-temporal graph whose underlying static graph (defined below) is acyclic is NP-hard and that this problem is polynomially solvable for contact-sequence graphs whose underlying static graph is acyclic. We remark in [13] that our proof of this is easily extended to show that finding foremost, fastest, min-hop, and shortest no-wait paths in interval-temporal graphs with an acyclic underlying static graph is NP-hard while these problems are polynomial for contact-sequence temporal graphs whose underlying static graph is acyclic.

The underlying static graph for any contact-sequence temporal graph is the graph that results when each edge $(u, v, t, \lambda)$ is replaced by the edge $(u, v)$ and then multiple occurrences of the same edge $(u, v)$ are replaced by a single edge $(u, v)$. For an interval temporal graph, its underlying static graph is obtained by replacing each edge ( $u, v$, intvls) by the edge $(u, v)$. Figure 5 shows the underlying static graphs for the temporal graphs of Figures 1 and 2. We show below that finding $m w f$ paths and walks and $m c f$ paths is NP-hard for the interval model but polynomially solvable for the contact-sequence model (for acyclic graphs).

Theorem 1. The mwf path and walk problems are NP-hard for interval-temporal graphs.
Proof. For the NP-hard proof, we use the sum of subsets problem, which is known to be NP-hard. In this problem, we are given $n$ non-negative integers $S=\left\{s_{1}, s_{2}, \ldots, s_{n}\right\}$ and another non-negative integer $M$. We are to determine if there is a subset of $S$ that sums to $M$. For any instance of the sum of subsets problem, we can construct, in polynomial time, the interval-temporal graph shown in Figure 6. For all edges other than $\left(u_{n}, v\right)$, the permissible departure times are from 0 through $M$ (i.e., their associated interval is $[0-M]$ ) and the edge $\left(u_{n}, v\right)$ has the single permissible departure time $M$ (equivalently, its associated interval is
[M-M] or simply $[M])$. The travel time for edge $\left(u_{i}, u_{i+1}\right)$ is $s_{i}$, that for $\left(u_{n}, v\right)$ is 1 , and that for the remaining edges is 0 . For every subset of $S$, there is a no-wait path from $u_{0}$ to $u_{n}$ that arrives at $u_{n}$ at a time equal to the sum of the $s_{i}$ s in that subset. Further, every no-wait path from $u_{0}$ to $v$ must get to $u_{n}$ at time $M$. Hence, there is a no-wait path from $u_{0}$ to $v$ if $S$ has a subset whose sum is $M$. In addition, every time-respecting path from $u_{0}$ to $v$ in the temporal graph of Figure 6 is a foremost path, as there is no path from $u_{0}$ to $v$ that can arrive at $v$ either before or after time $M+1$. Hence, a min-wait foremost path from $u_{0}$ to $v$ has a total wait of 0 if there is a subset of $S$ that sums to $M$; this path gets to $v$ at time $M+1$. Hence, the min-wait foremost path problem is NP-hard. The same construction shows that the min-wait foremost walk problem is also NP-hard for interval temporal graphs as every walk in the graph of Figure 6 is a path.


Figure 5. Underlying static graph for temporal graphs of Figures 1 and 2.


Figure 6. Interval-temporal graph for NP-hard proofs.
We note that the construction used in the above proof is easily modified, so that every edge has a travel time that is $>0$.

Theorem 2. The mcf path problem is NP-hard for interval-temporal graphs.
Proof. For this proof, we use the partition problem: Given $S=\left\{s_{1}, \ldots, s_{n}\right\}$ with $\left\{s_{1}+s_{2}+\right.$ $\left.\ldots+s_{n}=2 M\right\}$, we are to determine whether there is a subset whose sum is $M$; the $s_{i} \mathrm{~s}$ and $M$ are non-negative integers. We use the same graph as in Figure 6. Each edge ( $u_{i}, x_{i}$ ) has the cost $s_{i}$; all other edges have a cost of 0 . As in the proof of Theorem 1, every path from $u_{0}$ to $v$ corresponds to a subset of $S$; this subset consists of the $s_{i} \mathrm{~s}$ on the included edges of the form $\left(u_{i}, u_{i+1}\right)$. A path from $u_{0}$ to $v$ is feasible (time-respecting) if its length from $u_{0}$ to $u_{n}$ is $P \leq M$. Every feasible path from $u_{0}$ to $v$ gets to $v$ at $M+1$ and is a formost path. Feasible paths have the property that the sum, $P$, of the $s_{i} s$ in the associated subset is $\leq M$. Also, for every subset of $S$, there is a corresponding feasible path. The cost of such a path is
$2 M-P \geq M($ as $P \leq M)$. This cost takes on the min value $M$ if the sum of $s_{i} s$ on it is also $M$; i.e., if $S$ has a partition (i.e., a subset whose sum $P$ is $M$ ). Hence, the $m c f$ path problem is NP-hard for interval temporal graphs.

## 4. Min-Hop Foremost Paths

Before we develop the algorithm for finding mhf paths in interval-temporal graphs, we present the following theorems about mhf paths in temporal graphs.

Theorem 3. There exist interval-temporal graphs in which every mhf path from a source vertex s to a destination vertex $v$ has a prefix-path ending at a prefix vertex $u$ and the prefix-path from s to $u$ is not a min - hop path.

Proof. This can be seen from Figure 7. The only mhf path from $a$ to $d$ is $\langle a, 0, b, 1, c, 2, d\rangle$. However, the prefix path $\langle a, 0, b, 1, c\rangle$ is not a min-hop path from $a$ to $c$. The min-hop path from $a$ to $c$ is $<a, 8, c>$ which is not a prefix path to the only mhf path from $a$ to $d$, even though this mhf path goes via $c$.

Theorem 4. There exist interval-temporal graphs in which every mhf path from a source vertex s to a destination vertex $v$ has a prefix path ending at a prefix vertex $u$ and the prefix path from s to $u$ is not a foremost path.

Proof. This can also be seen from Figure 7. The only mhf path from $a$ to $f$ is $\langle a, 0, b, 7, d, 8, f\rangle$. However the prefix path $\langle a, 0, b, 7, d\rangle$ is not a foremost path from $a$ to $d$. There are two foremost paths from $a$ to $d$ in the interval-temporal graph of Figure 7, namely, $<a, 0, b, 1, c, 2, d>$ and $<a, 0, b, 1, c, 2, e, 3, d>$. Neither of these foremost paths from $a$ to $d$ is a prefix path to the only mhf path from $a$ to $f$, even though this $m h f$ path goes via $d$.

Theorem 5. There exist interval-temporal graphs in which every mhf path from a source vertex s to a destination vertex $v$ has a prefix path ending at a prefix vertex $u$ and the prefix path from s to $u$ is not an mhf path.

Proof. This can also be seen from Figure 7. The mhf path from $a$ to $f$ is $\langle a, 0, b, 7, d, 8, f\rangle$. However, the prefix path $\langle a, 0, b, 7, d\rangle$ is not an mhf path from $a$ to $d$. Out of the two foremost paths from $a$ to $d$ in the interval-temporal graph of Figure 7, namely, $<a, 0, b, 1, c, 2, d>$ and $\langle a, 0, b, 1, c, 2, e, 3, d\rangle$, the former has a fewer number of hops. Therefore, it is the the $m h f$ path from $a$ to $d$. However, it is not a prefix path for the only mhf path from $a$ to $f$, even though the only mhf path from $a$ to $f$ goes via $d$.

Theorem 6. Consider an interval graph $G$ that has a path from sto $v$. $G$ has an mhf path $P$ from $s$ to $v$ with the property that every prefix $Q$ of $P$ is an $h$ hop foremost path to $u$, where $h$ is the number of hops in $Q$ and $u$ is the last vertex of $Q$.

Proof. Since $G$ has a path from $s$ to $v$, it has a min-hop foremost path $R$ from $s$ to $v$. Let $S$ be the longest prefix of $R$ that is not an $h$-hop foremost path from $s$ to $u$, where $h$ is the number of hops in $S$ and $u$ is the last vertex in $S$. If there is no such $S$ then the theorem is proved. Assume $S$ exists. Replace $S$ in $R$, by an $h$-hop foremost path from $s$ to $u$, say $S^{\prime}$. The resulting path $R^{\prime}$ is an $m h f$ path from $s$ to $v$ with a prefix $S^{\prime}$ from $s$ to $u$ with an earlier arrival time at $u$, but the same number of hops. Repeating this replacement strategy a finite number of times, we obtain a min-hop foremost path $P$ from $s$ to $v$ that satisfies the theorem.


Figure 7. Example mhf paths from vertex $a$.

### 4.1. Algorithm to Find mhf Paths in Interval-Temporal Graphs

As noted in [13], some intervals of departure times on an edge may be redundant for the purposes of finding optimal foremost, min-hop, shortest, and fastest paths, as these are dominated by other intervals of the edge. These intervals are also redundant from the perspective of mhf paths (Theorem 6) and we assume that these redundant intervals have been removed in a preprocessing step.

Our algorithm employs the function $f(u, v, t)[9,13]$, which determines the earliest possible departure time $\geq t$ using the edge $(u, v)$.

### 4.1.1. Data Structures Used by mhf Algorithm 1

1. We use the same data structure to represent the interval temporal graph as we used in our earlier paper [13]. The data structure comprises a (say) C++ vector with one slot for each vertex in the graph. The slot for any vertex $u$ itself contains a vector of vertices adjacent from $u$. Associated with each adjacent vertex $v$ from $u$, there is a vector of time-ordered tuples for the edge $(u, v)$.
2. incSt is a structure that keeps track of vertices discovered in every hop. The fields in this structure are as follows:
(a) curVtxId is the current vertex.
(b) $\operatorname{arrTm}$ is the time of arrival at the current vertex.
(c) refProIncSt is reference to previous incSt that stores similar information about previous vertex on this path.
3. allHopPaths—array of lists that stores a list of vertices discovered at every hop. This array has, at most, $H$ lists, where $H$ is the maximum number of hops in min-hop paths from source vertex, $s$, to any of the vertices $v \in V$. Every element of the list is an instance of the structure incSt.
4. $\quad t_{E K A}$-array that stores for each vertex $v$ tuples of:
(a) earliest known arrival time $t_{f}$.
(b) number of hops in which earliest time found $h_{f}$.
(c) index, ind $x$, into allHopPaths to the structure incSt in the list at $h$.

### 4.1.2. Algorithm Description

- INPUT:

1. Temporal graph represented by data structure described in Section 4.1.1, item 1
2. Source vertex $s$

- OUTPUT:
- Array $t_{E K A}$ as described in Section 4.1.1, item 4

As an example, consider the interval-temporal graph of Figure 8. Let the source vertex be $S$ and $t_{\text {start }}=0$. In the first round (hopCnt $=1$ ), the neighbors $A, B$, and $C$ are identified as one-hop neighbors of $S$ with one-hop path arrival times of 1,5 , and 10 , respectively. Therefore, the earliest known arrival time to these neighbors are updated with $t_{E K A}$, with a hopCnt of 1 . In the next round (hopCnt $=2$ ), these one-hop paths are expanded to two-hop paths to vertices $B(S, A, B)$ and $C(S, B, C)$. The arrival times of these paths are 2 and 6 , which is earlier than their current arrival times. Therefore, their earliest known arrival times (foremost arrival times) are updated. In the third round (hopCnt $=3$ ), the earlier arriving 2-hop paths to $B$ and $C$ are expanded. While the 2-hop path to $C$ cannot be expanded any further, the 2-hop path to $B$ is expanded to obtsin a 3-hop path to $C$ that gets to $C$ at 4 , which is earlier than its current arrival time. Therefore, $t_{E K A}$ is again updated for this vertex. This path is expanded in the next round (hopCnt $=4$ ) and the 4 -hop path to $D$ $(S, A, B, C, D)$ is discovered, which is also the foremost arrival time at $D$. This path arrives at $D$ at 5 . The algorithm now terminates as hopCnt $=4=V-1$.


Figure 8. $m h f$ paths in interval-temporal graphs.

```
Algorithm 1 Min-hop foremost path algorithm
    Create startSt \(\leftarrow\left(s, t_{\text {start }}\right.\), null \()\) as an instance of incSt.
    Initialize \(t_{E K A}[s] \leftarrow t_{s t a r t} ; \forall v \in V, v \neq s, t_{E K A}[v] \leftarrow \infty\); Initialize allHopPaths \([0] \leftarrow\)
    \{startSt \(\}\)
    hopCnt \(=0 ;\) newVsInHop \(=1 ;\) tot \(V\) sRchd \(=1\)
    while (hopCnt \(<V-1\) ) and (newVsInHop \(\geq 0\) ) do
        hopCnt + +
        newVsInHop \(\leftarrow 0\)
        for each (refIncSt \(\in\) alHopPths \([\) hopCnt -1\(]\) ) do
            vert \(=\) refIncSt.curVtxId
            \(t_{\text {vertArr }}=\) refIncSt.arrTm
            for each (nbr \(\in V[v e r t] . n b r s)\) do
                \((\) depTm,intvlId \()=f\left((v e r t, n b r), t_{\text {vertArr }}\right)\)
                if \(\mathrm{depTm} \geq \infty\) then
                    continue \(\quad \triangleright\) start next loop iteration
                    end if
                    neww \(^{\prime 2 r} r_{n b r}=\operatorname{depTm}+\lambda_{\text {intolId }}\)
                if \(\left(\right.\) newArr \(\left._{n b r}<t_{E K A}[n b r] . t_{f}\right)\) then
                    if \(\left(h o p C n t>t_{E K A}[n b r] . h_{f}\right)\) then
                        Create an instance of incStruct as nis
                        nis \(\leftarrow\left(n b r, n e w A r r_{n b r}\right.\), refIncSt \()\)
                            \(t_{E K A}[n b r] . h_{f}=h o p C n t\)
                            alHopPths[hopCnt].append(nis)
                            \(t_{E K A}[n b r] . i n d x=n e w V s I n H o p++\)
                            else \(\quad \square\) arrival with an earlier time in same hop from a diferent prev node
                            Update
                            alHopPths \(\left[\right.\) hopCnt] \(\left[t_{E K A}[n b r]\right.\).ind \(\left.x\right]\)
                    end if
                    if \(t_{E K A}[n b r] . t_{f} \geq \infty\) then
                        totVsRchd ++
                    end if
                    \(t_{E K A}[n b r] . t_{f}=n e w A r r_{n b r}\)
                end if
            end for
        end for
    end while
```

Theorem 7. Algorithm 1 finds mhf paths from the source vertex, $s$, to all reachable vertices $v \in V$ in the interval-temporal graph $G=(V, E)$

Proof. The proof follows from Theorem 6 and the observation that the algorithm constructs mhf paths first with 1 hop, then with 2 hops, and so on. Form Theorem 6, it follows that, on each round, it is sufficient to examine only 1-hop extensions of paths constructed in the previous round.

The asymptotic complexity of Algorithm 1 is $O\left(N M_{i t g} \log \delta\right)$, where $N$ and $M_{i t g}$ are the number of vertices and edges, respectively, in the interval temporal graph and $\delta$ is the maximum number of departure intervals on an edge from the given $M_{i t g}$ edges. This is the same as that of the min-hop algorithm in our earlier paper [13].

## 5. Min-Wait Foremost Walks in Interval Temporal Graphs

### 5.1. Properties of mwf Walks

In this section, we describe properties of $m w f$ walks that are used later in the section for developing single-source all-destinations $m w f$ walks algorithm and the correctness proof of our $m w f$ algorithm. Some of the terminology we use is given below.

1. For a walk $X, t_{X}$ denotes its arrival time at its terminating vertex $v$ and $w_{X}$ denotes the total wait time accumulated by $X$ along its way from $s$ to $v$.
2. Comparing any two walks $X$ and $X^{\prime}$ w.r.t mwf arrival at a vertex $v, X$ is said to be the same or better than $X^{\prime}$ if $t_{X} \leq t_{X^{\prime}}$ and $w_{X} \leq w_{X^{\prime}}$.
3. A walk from $s$ to $v$ via some vertex $u$ is denoted as $W(s, u, v)$.
4. A walk $X$ from $s$ to $u$ that is extended further to obtain a walk $W(s, u, v)$ is a prefix walk denoted by $X=\operatorname{Pre}(W(s, u, v))$.

Definition 1. Walk dominance-for any two walks $A$ and $A^{\prime}$ from s to $u, A$ is said to dominate over $A^{\prime}$ if, for any walk $W(s, u, v)$ where $A^{\prime}=\operatorname{Pre}(W(s, u, v)), A^{\prime}$ can always be replaced by $A$ to produce a same-or-better walk $W^{\prime}(s, u, v)$ w.r.t mwf arrival at $v$

Theorem 8. If there are two walks, $A\left(t_{A}, w_{A}\right)$ and $A^{\prime}\left(t_{A^{\prime}}, w_{A^{\prime}}\right)$, that arrive at a vertex $u$, such that $t_{A} \leq t_{A^{\prime}}$, then if $\left(\left(t_{A^{\prime}}-t_{A}+w_{A}\right) \leq w_{A^{\prime}}\right)$ then $A$ dominates $A^{\prime}$.

Proof. We have:

$$
\begin{equation*}
\left(t_{A^{\prime}}-t_{A}+w_{A}\right) \leq w_{A^{\prime}} \tag{1}
\end{equation*}
$$

Let the departure time of $A^{\prime}$ for any onward walk from $u$ be $t_{\text {dep }} \geq t_{A^{\prime}}$. Since $t_{A} \leq t_{A^{\prime}}$, $t_{\text {dep }} \geq t_{A}$, therefore, $A$ can also depart vertex $u$ at $t_{\text {dep }}$. If $A^{\prime}$ departs at $t_{\text {dep }}$, the wait time accumulated by the time of departure is $w^{\prime}=w_{A^{\prime}}+t_{d e p}-t_{A^{\prime}}$. If $A$ departs at $t_{d e p}$, the wait time accumulated by the time of departure is $w=w_{A}+t_{\text {dep }}-t_{A}$. Due to Equation (1) $w^{\prime} \geq w$. Hence, if $A^{\prime}$ is replaced by $A$, then any possible extension of $A^{\prime}$ can still depart at the same time and would have accumulated an equal or less wait time by the time it departs from $u$. Therefore, $A^{\prime}$ can always be replaced by $A$ in any possible extension of $A^{\prime}$ to obtain a same-or-better walk.

Theorem 9. For identifying non-dominated walks among all the walks arriving at a vertex $u$, all walks may be arranged in an increasing order of arrival time at vertex $u$ and only adjacent walks need be compared for dominance.

Proof. If there are three walks $(A, B, C)$ terminating at $u$ with arrival times $\left(t_{A} \leq t_{B} \leq t_{C}\right)$, we need to show that it is sufficient to compare only the adjacent walks to find and retain the dominant walks terminating at $u$. This follows from the following two conditions:

1. If $B$ survives $A$ (is not dominated by $A$ ), then comparing $C$ with $B$ should have the same result as comparing $C$ with $A$. In other words
(a) If $C$ survives $B$, it also survives $A$.
(b) If $C$ is dominated by $B$, it is also dominated by $A$.
2. If $A$ dominates $B$ and $B$ dominates $C$, then $A$ also dominates $C$. Therefore, it does not matter in which order the walks are compared. Both $B$ and $C$ will be eliminated.
3. We are given $B$ is not dominated by $A$ or survives $A$. Based on the dominance criteria of Equation (1) we obtain Equation (2) for $B$ surviving $A$

$$
\begin{equation*}
\left(t_{B}-t_{A}+w_{A}\right)>w_{B} \tag{2}
\end{equation*}
$$

(a) When $C$ survives $B$, we have

$$
\begin{equation*}
\left(t_{C}-t_{B}+w_{B}\right)>w_{C} \tag{3}
\end{equation*}
$$

We are to prove that $C$ also survives $A$

$$
\begin{equation*}
\left(t_{C}-t_{A}+w_{A}\right)>w_{C} \tag{4}
\end{equation*}
$$

Adding Equations (2) and (3) yields (4)
(b) When $C$ is dominated by $B$, Equation (1) gives

$$
\begin{equation*}
\left(t_{C}-t_{B}+w_{B}\right) \leq w_{C} \tag{5}
\end{equation*}
$$

Replacing $w_{B}$ from Equation (2) in Equation (5), we obtain

$$
\begin{equation*}
\left(t_{C}-t_{A}+w_{A}\right) \leq w_{C} \tag{6}
\end{equation*}
$$

Therefore, $A$ also dominates $C$.
This proves that, when $B$ is not dominated by $A$, then comparing $C$ with $B$ has the same result as comparing $C$ with $A$.
2. Here, we prove that if $A$ dominates $B$ and $B$ dominates $C$, then $A$ also dominates $C$. Since $A$ dominates $B$ we have

$$
\begin{equation*}
\left(t_{B}-t_{A}+w_{A}\right) \leq w_{B} \tag{7}
\end{equation*}
$$

For $B$ dominates $C$ we have Equation (5). The two Equations (5) and (7), can be added together to obtain Equation (6), which states that $A$ dominates $C$. Therefore, when $A$ dominates $B$ and $B$ dominates $C$, regardless of the order in which the adjacent walks are compared, both $B$ and $C$ will be eliminated.
Hence, for any three walks $A, B$, and $C$ with $t_{A} \leq t_{B} \leq t_{C}$, only adjacent walks need to be compared to remove dominated walks and retain non-dominated walks. When applied transitively to all walks arriving at $u$, we see that if all the arriving walks are arranged in non-decreasing order of their arrival times, only adjacent walks need to be compared to eliminate the dominated walks and retain the non-dominated walks.

Theorem 10. For any two walks $A$ and $B$ arriving at a vertex $u$, where $t_{B}>t_{A}$, if $B$ is not dominated by $A$ as per the dominance criteria (1), then for any departure time $t$ available on an outgoing edge $(u, v)$ such that $t \geq t_{B}, A$ need not be considered for expansion.

Proof. Let $A^{\prime}$ be the walk obtained by extending $A$ from $u$ to a neighbor $v$ at time $t \geq t_{B}$. Similarly, $B^{\prime}$ is obtained by extending $B$ on $(u, v)$ at $t \geq t_{B}$. Assuming $\lambda$ as the travel time when departing at $t$ on the edge $(u, v)$, we have the following:

$$
\begin{align*}
& t_{A^{\prime}}=t+\lambda ; w_{A^{\prime}}=w_{A}+\left(t-t_{A}\right)  \tag{8}\\
& t_{B^{\prime}}=t+\lambda ; w_{B^{\prime}}=w_{B}+\left(t-t_{B}\right) \tag{9}
\end{align*}
$$

Therefore, we have $t_{A^{\prime}}=t_{B^{\prime}}$. In addition, since $B$ is not dominated by $A$, we have

$$
\begin{equation*}
t_{B}-t_{A}+w_{A}>w_{B} \tag{10}
\end{equation*}
$$

From Equations (8)-(10), we have $w_{B^{\prime}}<w_{A^{\prime}}$. This means that extending $A$ at $t \geq t_{B}$ is not beneficial, as it will be dominated by $B^{\prime}$.

Theorem 11. In order to find mwf walks, if walk $A$ arriving at $u$ is extended from $u$ to $v$ using an edge $(u, v)$ at time $t_{\text {dep }} \geq t_{A}$ in a departure interval $I(s, e, \lambda)$, then there is no benefit to extending A using the same edge $(u, v)$ in the same interval $I(s, e, \lambda)$ at a time $t^{\prime}$ where $t^{\prime}>t_{\text {dep }}$.

Proof. Let the extension of $A$ from $u$ to $v$ by departing at $t_{d e p}$ be $A 1$. Let the extension to $v$ obtained by departing at $t^{\prime}$ be $A 2$. We have the following two equations for the arrival and wait times of $A 1$ and $A 2$, respectively.

$$
\begin{gather*}
t_{A 1}=t_{\text {dep }}+\lambda ; w_{A 1}=w_{A}+\left(t_{\text {dep }}-t_{A}\right)  \tag{11}\\
t_{A 2}=t^{\prime}+\lambda ; w_{A 2}=w_{A}+\left(t^{\prime}-t_{A}\right) \tag{12}
\end{gather*}
$$

Clearly, $t_{A 1}<t_{A 2}$ as $t_{\text {dep }}<t^{\prime}$. Further, substituting $t_{A 2}, t_{A 1}, w_{A 1}$ from Equations (11) and (12) in expression $t_{A 2}-t_{A 1}+w_{A 1}$, we obtain

$$
\begin{equation*}
t_{A 2}-t_{A 1}+w_{A 1}=w_{A 2} \tag{13}
\end{equation*}
$$

From the dominance criterion of Equation (1), we conclude that $A 1$ dominates $A 2$. This means that if we extend $A$ at $t_{\text {dep }}$ and at $t^{\prime}>t_{\text {dep }}$ in the departure interval $I(s, e, \lambda)$ using the edge $(u, v)$ to obtain $A 1$ and $A 2$, respectively, $A 2$ will be dominated at $v$ by $A 1$. Therefore, there is no benefit of extending $A$ at $t^{\prime}$

Theorem 12. If the walk $A$ arriving at $u$ is extended from $u$ to $v$ using the edge $(u, v)$ at time $t_{\text {dep }} \geq t_{A}$ in a departure interval $I_{1}\left(s_{1}, e_{1}, \lambda_{1}\right)$, then $A$ may need to be extended again using the edge $(u, v)$ in a subsequent interval $I_{2}\left(s_{2}, e_{2}, \lambda_{2}\right)$ where $\left(s_{2}>e_{1}\right)$ if one of the following is true

1. $s_{2}+\lambda_{2}<t_{\text {dep }}+\lambda_{1}$
2. $\lambda_{2}>\lambda_{1}$

Proof. $t_{A} \leq e_{1}$ as $A$ can be extended in $I_{1}$. Therefore, the earliest time at which $A$ can be extended in $I_{2}$ is $s_{2}$. It need not be extended at any other time in $I_{2}$ due to Theorem 11. Let the extension of $A$ obtained by extending in $I_{1}$ be $A 1$ and that obtained by extending in $I_{2}$ be $A 2$.

1. If $s_{2}+\lambda_{2}<t_{\text {dep }}+\lambda_{1}$, it means $t_{A 2}<t_{A 1}$; therefore, $A$ will need to be extended at $s_{2}$ so we do not miss any opportunities of further expansion in intervals of departure available at $v$ at a time $t$, such that $s_{2}+\lambda_{2} \leq t<t_{\text {dep }}+\lambda_{1}$.
2. We have the following two equations for the arrival and wait times of $A 1$ and $A 2$, respectively

$$
\begin{gather*}
t_{A 1}=t_{d e p}+\lambda_{1} ; w_{A 1}=w_{A}+\left(t_{d e p}-t_{A}\right)  \tag{14}\\
t_{A 2}=s_{2}+\lambda_{2} ; w_{A 2}=w_{A}+\left(s_{2}-t_{A}\right) \tag{15}
\end{gather*}
$$

Comparing $A 1$ and $A 2$ at $v$, we see that $s_{2}+\lambda_{2} \geq t_{\text {dep }}+\lambda_{1}$ as otherwise, condition 1 would be true. Therefore, we have:

$$
\begin{equation*}
t_{A 2}-t_{A 1}=s_{2}+\lambda_{2}-t_{d e p}-\lambda_{1} \tag{16}
\end{equation*}
$$

Let $e=t_{A 2}-t_{A 1}+w_{A 1}$. Evaluating $e$ using Equation (16) and substituting for $w_{A 1}$ from Equation (14), we obtain

$$
\begin{equation*}
e=w_{A}+s_{2}-t_{A}+\left(\lambda_{2}-\lambda_{1}\right) \tag{17}
\end{equation*}
$$

Substituting $w_{A 2}$ from Equation (15) into Equation (17), we obtain:

$$
\begin{equation*}
e=w_{A 2}+\left(\lambda_{2}-\lambda_{1}\right) \tag{18}
\end{equation*}
$$

We know that for $A 2$ to survive being dominated by $A 1$, we need Equation (19) to be true

$$
\begin{equation*}
e>w_{A 2} \Longrightarrow w_{A 2}+\left(\lambda_{2}-\lambda_{1}\right)>w_{A 2} \tag{19}
\end{equation*}
$$

Therefore, for Equation (19) to be true, $\lambda 2>\lambda 1$, otherwise $A 2$ will be dominated by $A 1$.
Therefore, unless 1 or 2 is true, there is no benefit of extending in interval $I_{2}$.
This can also be seen from the example of Figure 9. There are two intervals on the edge $(u, v)$ with $I 1(2,6,12)$ and $I 2(10,20,5)$. A departs at 2 and then does not need to depart in I2 $(10,20,5)$, as none of the conditions of Theorem 12 is satisfied. $A^{\prime}$ departs at 4 in $I 1(2,6,12)$ but needs to depart in $I 2(10,20,5)$ again as condition 1 is satisfied.


Figure 9. Example of departure in next intervals.
Theorem 13. If walk $A$ arriving at $u$ is extended from $u$ to $v$ using the edge $(u, v)$, it should always be extended at $\left(t_{\text {dep }}\right.$, intvlId $)=f\left(u, v, t_{A}\right)$, where $f(u, v, t)$ is the next function referred to in Section 4.1 and described in [9,13].

Proof. From Theorem 11, a given walk departing in a departure interval $I$ on an edge $(u, v)$ should always depart at the earliest possible departure time in I. Further, as per the conditions stated in Theorem 12, it is evident that, for any two available departure intervals $I_{1}\left(s_{1}, e_{1}, \lambda_{1}\right)$ and $I_{2}\left(s_{2}, e_{2}, \lambda_{2}\right)$ where $e_{1} \leq s_{2}$ on an edge $(u, v)$, a walk $A$ terminating at $u$ such that $t_{A} \leq e_{1}$ should always depart in $I_{1}$ and whether or not it departs in $I_{2}$ is determined by the conditions specified in the theorem.

### 5.2. Departure from Source Vertex s

All walks that start from a source vertex $s$ have a wait time of 0 , as there is no wait accumulated at the source vertex. Therefore, for any available departure interval on an edge $(s, x)$, we should be extending at every possible instance of time in this departure interval. Theorems 11 and 12 do not apply to walks departing from the source vertex $s$, as these theorems assume that there is an extra wait accumulated if a walk $A$ is extended at a later time from the departing vertex $u$. However, this is not true when $u=s$ as the wait time at $s$ is always 0 . This implies, we may assume that $m w f$ walks do not have a cycle that involves $s$ as such cycles can be removed from the walk without increasing either the total wait time or the arrival time at the destination. From the source vertex $s$, walks may depart at every departure instance available to find $m w f$ walks. To account for this, we introduce the concept of a walk class.

Definition 2. A walk class $(w s, w e, u)$ is a set of walks that arrives at the vertex $u$ with a wait time of 0 . The first walk in this set arrives at ws and the last one arrives at we, where we $>$ ws. There is a walk in the walk class at every instance of time in the range $[w s, w e]$ and each of these walks has the same total wait time, which is 0 . This can be seen in Figure 10.


Figure 10. Example of a walk class.
For every departure interval denoted by $I(s t a r t, e n d, \lambda)$ where $e n d>$ start on an edge $(s, x)$, we need to generate a walk class, ranging from the start to end, with travel time as $\lambda$ and a wait time of 0 . Once the walks in the walk class arrive at the neighbor $x$, Theorems 11 and 12 apply to each walk in the class.

Theorem 14. Each walk in a walk class survives the walk, if any, that arrives before it.

Proof. This is easy to see from Equation (1), as these walks have the same wait time but different arrival times.

Theorem 15. If two walk classes $W C_{1}$ and $W C_{2}$ arrive at a vertex $u$, there is no dominance to be considered.

Proof. This follows from Equation (1), as the wait times of all walks is 0 . When they overlap, one of them may be discarded. When they do not overlap, both can be retained.

When a walk class, $W C(w s, w e)$, terminates at a vertex $u$, if some portion of it overlaps with a departure interval $I(s, e, \lambda)$ from $u$ to a neighbor $v$, the set of walk instances $x$ in the walk class $W C$ such that $s \leq t_{x} \leq e$ can be immediately extended to the neighbor $v$ without any additional wait at $u$ into a new walk class arriving at $v$ as $W C^{\prime}\left(w s^{\prime}, w e^{\prime}\right)$. Therefore, each walk in $W C^{\prime}\left(w s^{\prime}, w e^{\prime}\right)$ also has a wait time of 0 . If there is no overlap with a departure interval from $u$ to $v$, then only the walk instance at we needs to be extended to $v$ due to Theorems 10 and 14.

For the $m w f$ problem, redundant intervals as described in [13] for the foremost, minhop, fastest problems are not redundant and need to be retained. Such intervals, however, are redundant for mhf problems, as noted in Section 4.1

This is illustrated with the example of Figure 11. The min-wait foremost $m w f$ walk from $s$ to $c$ would benefit by departing from the vertex $s$ at 0 , instead of waiting for the next faster interval that becomes available for departure at 2 and reaches $a$ at 4 . Departing at 0 from $s$ reaches $a$ at 5. In this case, the $m w f$ walk would be $W_{m w f}=(s, 0, a, 7, b, 1, c)$ arriving at vertex $c$ at time 9 with a wait time of 0 .


Figure 11. Interval temporal graph with some slow intervals.

### 5.3. Additional Data Structures

For finding $m w f$ walks, we introduce an additional data structure which is a sorted list of all departure intervals in the interval temporal graph, sorted by the arrival time when departure is at the start of the interval. Each interval in this list is represented as $I(s, e, \lambda)$ and the sort key for the list is $s+\lambda$. This is demonstrated in Figure 12 for the interval temporal graph of Figure 11.


Figure 12. Interval list in non-decreasing order of arrival.

### 5.4. Algorithm to Find mwf Walks

5.4.1. Data Structures Used by mwf Algorithm 2

1. Input graph $G$ as in our earlier paper [13] and briefly described in Section 4.1.
2. intvlsInfo is a structure that describes a departure interval.
(a) $u$ is the start vertex.
(b) $\quad v$ is the end vertex.
(c) st-start time of the interval
(d) end-end time of the interval.
(e) $\lambda$ is the travel time on this intvl.
(f) intvlid is the Id of this intvl on connection $(u, v)$ in the input graph
3. intelList—array of intvlInfo that contains all the departure intervals from the input graph, as described in Section 5.3
4. $P Q$ of ad-hoc intervals-Priority queue of items of type intvlsInfo. The priority key is (intvlStrt $+\lambda$ ).
5. mwfWClass is a structure that describes a walk class
(a) strtTm—arrival time of first walk instance in this class
(b) endTm-arrival time of last walk instance
(c) $w t T m$-total wait time accumulated on the walk.
(d) lastExpAt-This is an array of $\lambda$ s of the departure interval in which this walk was last expanded to each of the nbr of its terminating vertex. The number of items in this array is the number of nbrs of the vertex at which this walk terminates.
6. listWClasses-This is an array of lists of walk classes. The size of the array is the number of vertices in the graph. Each list in the array is the list of walk classes terminating at the vertex corresponding to the array index. This list is always sorted in the increasing order of arrvTmStrt of the walk classes.
7. $m w f W a l k s-A n ~ a r r a y ~ o f ~ w a l k ~ c l a s s e s . ~ A r r a y ~ h a s ~ a n ~ i t e m ~ f o r ~ e a c h ~ v e r t e x ~ i n ~ t h e ~ g r a p h . ~$ For each vertex, it has the $m w f$ walk to that vertex.

### 5.4.2. Algorithm Description

- INPUT:

1. Temporal graph represented by data structure described in Section 5.4.1, item 1
2. Source vertex $s$

- OUTPUT:
- Array mwfWalks as described in Section 5.4.1, item 7

For simplicity, we assume that all the edge travel times are $>0$. Our algorithm for solving the $m w f$ problem (Algorithm 2) can easily be extended to the case when some travel times are $\geq 0$. Later, we show how to carry out this extension. Algorithm 2 looks at all possible departure intervals in non-decreasing order of their arrival times when departing at the start of the interval (intvl.strtTm + intvl. $\lambda$ ). In Step 8, it saves the arrival time of the latest interval being considered in the variable curTm. In the subsequent while loop, it keeps fetching new intervals as long as they have the same arrival time as curTm, when departing at the start of the interval. For each fetched interval newintvl, at the originating vertex $u$ of newintvl, it finds a walk class prevW with latest arrival time of the start of the walk class prevW.strtTm such that prevW.strtTm $\leq$ newintvl.st. No other walk arriving at $u$ needs to be considered for expansion on this interval due to Theorem 10. Using this previous walk at the originating vertex $u$, a new walk is created from $u$ to $v$ at the start of the newIntvl using the function createNewW. This function creates a new walk from the previous walk departing at start of newIntvl only if the second part of Theorem 12 is satisfied. The first part of Theorem 12 is already taken care of because the intervals are examined only in non-decreasing order of intvl.st + intvl. $\lambda$.

```
Algorithm 2 Min-wait foremost walks algorithm
    Create a new newWClass \(\leftarrow\left(t_{\text {start }}, \infty, 0, n b r s \_s[0]\right)\) as an instance of \(m w f W C l a s s\). Val-
    ues are assigned to fields arrivTmStrt, arrivTmEnd, wtTm, lastExpAt, respectively. Last
    field is an array of size out degree of vertex \(s\).
    Initialize mwfWalks[s] newWClass; initialize
    listWClasses[s].pushback(newWClass); \(\forall v \in V, v \neq s, m w f W a l k s[v] \leftarrow\) null
    nodesRchd \(\leftarrow 1\); staticIntvls \(\leftarrow\) intvlList.size; indxStaticIntvls \(\leftarrow 0\)
    setupNewWClass(s,newWClass.arrivTmStrt)
    reachableNodes \(\leftarrow \operatorname{get}\) ReachableNodes \((G, s)\)
    newIntvl \(\leftarrow\) removeMinIntvl ()
    while ((indxStaticIntvls \(<\) staticIntvls) or !PQ.empty()) and (nodesRchd \(<\)
    reachableNodes) do
        curTm \(\leftarrow\) newIntvl.st + newIntvl. \(\lambda\)
        \(u=\) newIntvl.u; \(v=\) newIntvl.v
        newVs \(\leftarrow 0\)
        while (curTm \(==\) newIntvl.st + newIntvl. \(\lambda\) ) do
            if (newIntvl. \(v \neq s\) ) then
                prevWIndx \(\leftarrow\) getPrveW(newIntvl)
                if prevWInd \(x \neq-1\) then
                        newW \(\leftarrow\) createNewW \((\) listWClasses \([u][\) prevWIndx], newIntvl \()\)
                    if newW then
                            ins \(\leftarrow\) insNewW \((\) new \(W, v) \quad \triangleright\) Inserts new \(W\) in list of walks at \(v\). If
    new \(W\) is best \(m w f\) walk it is recorded in mwfWalks \([v]\)
                    if ins \(==N E W\) then
                        newVs ++
                    end if
                    if ins \(\neq 0\) then
                                    setupNewWClass(v,newW.strtTm)
                    end if
                end if
                end if
            end if
            newIntvl \(\leftarrow\) removeMinIntvl ()
        end while
        nodesRchd \(+=\) newVs
    end while
```

After the new walk is created on an edge $(u, v)$, it is appended to the list of walks at $v$. To qualify as a valid walk to be appended to the list of walks at $v$, it needs to be compared for dominance with only the last walk in the list at $v$ due to Theorem 9. This walk is appended to the list at $v$ only if it is not dominated. The insert function also checks if this is the first walk arriving at $v$. If so, it updates the $m w f$ Walks for vertex $v$ and increments the count of newVs. If this is not the first walk to arrive at $v$, but if it is better w.r.t mwf arrival than the best known walk at $v, m w f$ Walks is updated for vertex $v$. Note that the new walk cannot have an earlier arrival time than the best known walk at $v$ as the intervals are examined in non-decreasing order of arrival times, but it may have a lesser wait time.

After a walk has been appended to the list of walks at $v$, the function setupNewWClass examines all the neighbors of $v$. If the start time of the last arriving walk class is in the middle of an available departure interval, intvl given, in the input graph $G$ on an edge $(v, w)$, a new departure sub-interval of intvl, say subIntvl, is created. The start time of subIntvl is the arrival time of the walk and travel time same as that of the intol (intol. $\lambda$ ). This way, every walk originates at the start of an interval given in graph $G$ or at start of a subIntel that is a sub-interval of one of the intervals given in the input graph G. This new subIntvl is inserted into the priority queue $P Q$. In the removeMinIntvl function, the
interval with minimum arrival time (when departing at the start of the interval) from intvlList and the top of $P Q$ is returned.

Theorem 16. Algorithm 2 finds mwf walks from s to all other vertices in an interval temporal graph when all edge travel times are $>0$.

Proof. The algorithm examines all the departure intervals in non-decreasing order of arrival time, when departing at the start of the interval. For every interval, newintvl in that order, it considers the best walk eligible for extension in this interval, as per Theorem 10, by obtaining the index for the previous walk (prevWIndex) in the list of walks at the departure vertex $u$, in Step 13. It extends the prevWalk with departure in newIntvl only if this extension is beneficial as per Theorem 12. Once the walk is extended, it updates the lastExpAt for prevWalk for the edge $(u, v)$, so that subsequent departure intervals on $(u, v)$ are used for expanding prevWalk only if the resultant walks may be beneficial, as per Theorem 12.

The extended walk is appended to the list of walks at $v$ only if it is not dominated as per the dominance criterion of Equation (1). In addition, $m w f$ walks at $v$ are updated if the arriving walk is better than the previous $m w f$ walk, w.r.t $m w f$ arrival at $v$ or if it is the first walk to arrive at $v$.

Finally, the arriving walk at $v$, new $W$, is examined w.r.t all outgoing edges $(v, w)$. If for a departure interval $I(s, e, \lambda)$ to any neighbor $w$ we have $I_{s}<n e w W . s t r t T m \leq I_{e}$, a new sub-interval, subIntol, is created that has a start time of newW.strtTm and inserted into $P Q$. When such a subIntvl is the minimum interval among all the intervals from the input graph $G$ and the subIntvls created during the course of Algorithm 2, the newW gets the opportunity to expand at the start time of subIntvl.

Therefore, this algorithm starts at the start vertex and generates walks at all feasible departure instances w.r.t $m w f$ walks criteria. Feasible departure intervals are examined in non-decreasing order of arrival times; therefore, every new walk generated has an arrival time $\geq$ arrival time of any previously generated walk. No feasible non-dominated walk is missed due to Theorems 8-13.

Every non-dominated walk generated is preserved and expanded as per Theorems 11-13. When a walk terminating at a vertex $v$ is discovered that is the best among the walks arriving at $v$ w.r.t $m w f$ arrival, it is recorded as the $m w f$ walk at $v$. The algorithm terminates only when one of the following is true

1. All intervals including the sub - intervals created during the course of Algorithm 2 have been examined.
2. mwf walks to all reachable vertices have been found and the arrival time when departing at the start of next interval being examined is bigger than the max arrival time of the $m w f$ walks found. Therefore, examining any more intervals cannot give a better $m w f$ walk.

Therefore, when the algorithm terminates, $m w f$ walks to all reachable vertices have been found.

For handling the case when some travel times are 0 , in the while loop of line 11, we need to first obtain all the edges with 0 travel times arriving at currTm and form their transitive closure. Then, we can process every edge arriving at the same currTm over this transitive closure.

Theorem 17. When time is an integer, Algorithm 2 has pesudopolynomial complexity.
Proof. Let $T$ be the maximum possible arrival time at any vertex. Since $c u r T m$ is an integer that increases in each iteration of the outer while loop, this outer loop is iterated $O(T)$ times. For each value of cur Tm , the while loop of line 11 is iterated $O\left(M_{i t g} T\right)$ times, where $M_{i t g}$ is number of edges in the input interval temporal graph, as the number of walks that
can arrive at a vertex $v$ at time curTm is $O$ (indegree $(v) * T)$. The components of the time required for each iteration of the while loop of line 11 are

1. $\log ($ walks $(v))$ in step 13 to find the previous walk that can be extended at the current departure instance $\left(t_{\text {arr }}\right)$. walks $(v)$, which is the number of walks at $v$ at time $t_{\text {arr }}$, can be at most $t_{\text {arr }}$ as each preserved walk at a given vertex $v$ has a unique arrival time. Therefore, time taken by this step is $\log \left(t_{\text {arr }}\right)$ or at most $\log (T)$.
2. Constant time for creating a new walk newW in step 15.
3. Constant time for inserting new $W$, in the list of walks $l$ at $v$ in step 17 . $l$ at every vertex $v$ maintains a list of walks that has arrived at $v$ in non-decreasing order of their arrival times. new $W$ only needs to be compared with the last walk in $l$.
4. At most $d \log (\delta)$ in step $22, d$ is the out-degree of the vertex $v$ at which newW terminates and $\delta$ is the number of departure intervals on an outgoing edge $(v, w)$ that has maximum departure intervals among all the outgoing edges from $v$. Each of the departure intervals on outgoing edges from $v$ may need to be added to $P Q$.
Size of $P Q$ can be, at most, the maximum number of departure instances in the graph $G$, which is $M_{i t g} T$. Therefore, the maximum time taken in this step is $d \log (\delta) \log \left(M_{i t g} T\right)$.
5. Finally, in step 27, departure instance with minimum arrival time is retrieved. Maximum size of $P Q$ can be $M_{i t g} T$. Therefore, maximum time taken by this step is $\log \left(M_{i t g} T\right)$
Therefore, the complexity of the algorithm is $O\left(T\left(M_{i t g} T(\log (T)+d \log (\delta)\right.\right.$ $\left.\left.\left.\log \left(M_{i t g} T\right)+\log \left(M_{i t g} T\right)\right)\right)\right)$, which is $O\left(M_{i t g} T^{2} \log (T)\right)$. Hence, the time complexity of Algorithm 2 is pseudopolynomial.

## 6. Linear Combination of Optimization Criteria

Bentert et al. [10] present a polynomial time algorithm that optimizes any linear combination of the eight optimization criteria foremost, reverse_foremost, fastest, shortest, min_hop, cheapest, most_likely, and min_wait in a contact-sequence temporal graph. In this section, we show that, when time is discrete, we can use the algorithm of Bentert et al. [10] to solve the $m h f, m w f$ and $m c f$ problems in polynomial time for contact-sequence graphs. This algorithm may also be used to find $m h f, m w f$, and $m c f$ paths in interval temporal graphs by first transforming the interval-temporal graph into an equivalent contact sequence graph as described in Section 1. Since this transformation may increase the number of edges by an exponential amount, this approach does npt result in a polynomial time algorithm for interval-temporal graphs. However, it does result in a pseudopolynomial time algorithm as shown in Theorem 17.

To solve $m h f$, for example, for contact-sequence graphs, we perform the following:

1. Set the coefficient for every criterion other than min_hop and foremost to 0 .
2. The coefficient for the foremost criterion may be set to any integer that is greater than or equal to the number of vertices, $n$, in the contact-sequence graph.
3. Set the coefficient for the min-hop criterion to 1.

With these settings, Bentert et al. [10] will find walks that minimize $h(s, v)=c_{f} *$ $t_{a}(v)+\operatorname{hops}(v)$, where $c_{f}$ is the coefficient for the foremost criterion, $t_{a}(v)$ is the time at which the walk from $s$ arrives at $v$, and $\operatorname{hops}(v)$ is the number of hops in the walk. Since a min-hop walk is necessarily a min-hop path, it has no more than $n-1$ hops. If we examine the digits of $\operatorname{mhf}(s, v)$, which is a non-negative integer, using the radix $c_{f}$, the least significant digit is the number of hops and the remaining digits give $t_{a}$. Hence, $\operatorname{mhf}(s, v)$ is minimized by min-hop foremost paths to $v$ and not by any other path or walk.

The strategy for $m w f$ and $m c f$ is similar. In the case of $m w f$, the function to optimize is $m w f(s, v)=c_{f} * t_{a}(v)+w a i t(s, v)$ and for $w c f$, the optimization function is $w c f(s, v)=$ $c_{f} * t_{a}(v)+\operatorname{cost}(s, v)$, where wait $(s, v)$ is the total wait time on the walk from $s$ to $v$ and $\operatorname{cost}(s, v)$ is its cost. For $m w f(s, v), c_{f}$ must be larger than the maximum total wait time on an optimal foremost walk from $s$ (a simple bound is the maximum arrival time of a foremost path to reachable vertices), and for $m c f(s, v), c_{f}$ must be larger than the maximum
cost of an optimal foremost path (a simple bound to use is the sum of the maximum cost of each edge). For simplicity, we use $c_{f}=2^{32}$ in our experiments, as this is large enough for our data sets.

It is easy to see how the above modeling strategy may be used to find, say, min-cost foremost paths with the fewest number of hops.

## 7. Experimental Results

In this section, we compare the performance of our mhf path and $m w f$ walk algorithms to the algorithm of Bentert et al. [10]. Since the latter algorithm works only on contactsequence graphs, we first transform our interval-temporal graphs into equivalent contactsequence graphs (as noted in Section 1, this can be performed when time is discrete) and then we use the strategy discussed in Section 6 to formulate an appropriate optimization function for the algorithm of Bentert et al. [10]. Since the optimization function constructed in Section 6 has values larger than what can be handled by the datatype int, we modified the code of Bentert et al. [10] using the datatype long for relevant variables.

Our experimental platform was an IntelCore, $i 9-7900$ XCPU@3.30GHz processor with 64 GB RAM. The C++ codes for finding the optimal linear combination of multiple optimization criteria for contact-sequence temporal graph algorithms was obtained from the authors of [10]. All other algorithms were coded by us in C++. The codes were compiled using the $g++$ ver.7.5.0 compiler with option O2. For test data, we used the datasets used in $[8,10]$, for the contact-sequence graphs. We transformed these contact-sequence graphs to interval-temporal graphs to run on our algorithms. We also generated some synthetic datasets as interval-temporal graphs that we transformed to contact-sequence models so the programs of Bentert et al. [10] could be run on them.

### 7.1. Datasets

We used 13 of the 14 real-world contact-sequence graphs that were used in [8,10,12,13]. The 14th dataset, $d b l p$, was not used as it had a few negative timestamps. The statistics for the remaining 13 real-world datasets are given in Table 1. In this table, $|V|$ is the number of vertices, $\left|E_{S}\right|$ is the number of edges in the underlying static graph and cs - edges is the number of edges in the contact-sequence temporal graph. The ratio (cs-edges) / $\left|E_{s}\right|$ is the temporal activity of the graph. Note that the number of edges in the interval-temporal graph is also $\left|E_{s}\right|$. The datasets have a wide range of sizes in terms of the number of vertices and edges. Consistent with $[8,10,12,13$ ], the travel time, $\lambda$, on all edges was set to 1 . The temporal activity on these datasets is low, ranging from 1 to 3.67.

Table 1. Koblenz collection graph statistics.

| Dataset | $\|\boldsymbol{V}\|$ | $\left\|\boldsymbol{E}_{\boldsymbol{s}}\right\|$ | $\boldsymbol{c s}-\boldsymbol{e d g e s}$ | Activity |
| :---: | :---: | :---: | :---: | :---: |
| epin | 131.8 K | 840.8 K | 841.3 K | 1 |
| elec | 7119 K | 103.6 K | 103.6 K | 1 |
| fb | 63.7 K | 817 K | 817 K | 1 |
| flickr | 2302.9 K | $33,140 \mathrm{~K}$ | $33,140 \mathrm{~K}$ | 1 |
| growth | 1870.7 K | $39,953 \mathrm{~K}$ | $39,953 \mathrm{~K}$ | 1 |
| youtube | 3223 K | 9375 K | 9375 K | 1 |
| digg | 30.3 K | 85.2 K | 87.6 K | 1.02 |
| slash | 51 K | 130.3 K | 140.7 K | 1.07 |
| conflict | 118 K | 2027.8 K | 2917.7 K | 1.43 |
| arxiv | 28 K | 3148 K | 4596 K | 1.45 |
| wiki-en-edit | $42,640 \mathrm{~K}$ | $255,709 \mathrm{~K}$ | $572,591 \mathrm{~K}$ | 2.23 |
| enron | 87,274 | 320.1 K | 1148 K | 3.58 |
| delicious | 4512 K | $81,988 \mathrm{~K}$ | $301,186 \mathrm{~K}$ | 3.67 |

In [13], we generated five synthetic datasets with higher activity and variable $\lambda$ s by starting with the social network graphs of youtube, flickr, livejournal shared by the authors
of [17] at http:/ / socialnetworks.mpi-sws.org/data-imc2007.html (accessed on 16 August 2022). These graphs represent user-to-user interactions.

Table 2 shows the statistics for the five synthetic temporal graphs generated by us.
Table 2. Synthetic graphs statistics.

| Graphs with $\mu_{I}=4, \mu_{\boldsymbol{D}}=\mathbf{5}, \boldsymbol{\mu}_{\boldsymbol{T}}=\mathbf{3}$ |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: |
| Dataset | $\|V\|$ | $\left\|E_{S}\right\|$ | cs-edges | Edge Activity |
| youtube | 1157.8 K | 4945 K | $105,039 \mathrm{~K}$ | 21.2 |
| flickr | 1861 K | $22,613.9 \mathrm{~K}$ | $480,172 \mathrm{~K}$ | 21.24 |
| livejournal | 5284 K | $77,402.6 \mathrm{~K}$ | $1,643,438 \mathrm{~K}$ | 21.3 |
| Graphs with $\mu_{I}=4, \mu_{D}=8, \mu_{T}=3$ |  |  |  |  |
| youtube | 1157.8 K | 4945 K | $159,103.7 \mathrm{~K}$ | 32.1 |
| flickr | 1861 K | $22,613.9 \mathrm{~K}$ | $727,405.9 \mathrm{~K}$ | 32.1 |

Table 3 gives the time (in seconds) required to read each dataset from the disk as well as the disk memory required by each dataset. The columns labeled [10] are for the case when the dataset is stored in the input format used by the algorithm of Bentert et al. [10] and those labeled Ours are for the input format required by our algorithm. As can be seen, the disk reading time and disk space required by the interval-temporal graph representation are less than that required by the Bentert et al. [10]. Further, this difference increases as the temporal activity increases. In fact, for four of the instances, input creation code of [10] failed to create the required input format for lack of sufficient memory (out-of-mem).

Table 3. Reading times and sizes.

|  | Koblenz Collection |  |  |  |
| :---: | :---: | :---: | :---: | :---: |
| Dataset | Reading Time (in s) |  | Sizes in MBs |  |
|  | $[10]$ | Ours | [10] Graph | Ours |
| epin | 0.46 | 0.34 | 31.1 | 19.3 |
| elec | 0.13 | 0.07 | 4.5 | 3.3 |
| fb | 0.58 | 0.34 | 32.9 | 21.5 |
| flickr | 19.3 | 14.2 | 1492.3 | 911 |
| growth | 27 | 19.9 | 2188.3 | 1485.4 |
| youtube | 6.1 | 4.2 | 456.1 | 300.8 |
| digg | 0.11 | 0.06 | 3.6 | 2.5 |
| slash | 0.14 | 0.07 | 7 | 5 |
| conflict | 1.44 | 0.82 | 98.8 | 43.2 |
| arxiv | 2.5 | 1.4 | 193 | 104 |
| wiki-en-edit | - | 227 | out-of-mem | 20,115 |
| enron | 0.74 | 0.36 | 55.9 | 29.3 |
| delicious | 209 | 78 | 16,249 | 7552 |
|  | Synthetic Datasets with $\mu_{I}=4, \mu_{D}=5, \mu_{T}=3$ |  |  |  |
| youtube | 67 | 4 | 4400 | 272.2 |
| flickr | - | 17.8787 | out-of-mem | 1248 |
| livejournal | - | 56.1638 | out-of-mem | 4411 |
|  | Synthetic Datasets with $\mu_{I}=4, \mu_{D}=8, \mu_{T}=3$ |  |  |  |
| youtube | 102 | 3 | 6775 | 272 |
| flickr | - | 18 | out-of-mem | 1249 |

As an example, for the large synthetic youtube graph with ( $\mu_{I}=4, \mu_{D}=8, \mu_{T}=3$ ) which has an activity factor of 32.1 , the reading time of the interval-temporal graph is 3.53
s while the program by Bentert et al. [10] takes about 102.5 s to read the corresponding graph in Bentert's transformed format. The size of the interval temporal graph for the same dataset on disk is 272.3 MB as compared to 6.7 GB for the Bentert's transformed graph from the corresponding contact-sequence graph.

### 7.2. Run Times

As in $[8,10,12,13]$, we assume that the the graph is resident in memory (i.e., the read time from disk is not accounted for). This is a valid assumption in applications where the graph is input once and queried many times. To use the algorithm of Bentert et al. [10], we use the following steps:

1. Transform the interval temporal graph to an equivalent contact sequence graph (csg) as described in Section 1.
2. Use Bentert's graph transformation program [10] to convert the $\operatorname{cs} g$ to the input format used by Bentert's linear combination algorithm.
3. Run Bentert's linear combination program on the transformed graph from Step 2 using the coefficients given in Section 6.
The step 2 transformation was unsuccessful on the Koblenz dataset wiki -en - edit as well as on all our synthetic datasets other than youtube graph with ( $\mu_{I}=4, \mu_{D}=5,8$, $\mu_{T}=3$ ), as indicated in Table 3. The failure of this step resulted from the unavailability of sufficient memory to run the step 2 transformation code of [10] on these instances. We measured the average of the runtimes for 100 randomly selected source vertices for each dataset except the delicious from the Koblenz collection and the synthetic datasets, where we used only five randomly selected source vertices. This reduction in the number of source vertices was necessary because of the excessive runtime taken by the algorithm of [10] on these datasets. The average run-times (in seconds) for the Koblenz and synthetic datasets are given in Table 4. This time does not include the time required by steps 1 and 2. The speedups (time taken by the algorithm of [10]/time taken by our algorithm) are also shown visually in Figures 13-16.

The speedups obtained by our $m h f$ algorithm over [10] range from 3 to 207.5 for the Koblenz datasets and from 451.3 to 679.56 for the synthetic datasets. Our mhf algorithm outperforms that of [10] on all datasets. The speedups obtained by our mwf algorithm over [10] range from a low of 0.045 to a high of 23.3. The algorithm of [10] outperforms our $m w f$ algorithm when the graph has very low connectivity. Their algorithm is able to quickly discover that no more vertices are reachable from the source vertex and, so, terminates sooner than our mwf algorithm. However, when there are many reachable vertices, our algorithm outperforms that of Bentert et al. [10]. We also note that our algorithm works on interval-temporal graphs where time intervals may be continuous or have large durations, while the algorithm of Bentert et al. [10] is only for contact-sequence graphs, which are a subset of interval temporal graphs. As noted earlier, the algorithm of [10] could not be run on some datasets because of the failure of step 2 as it ran out of memory. On the four datasets where the algorithm of [10] could not be run, the average runtime of our mhf algorithm was 1, 1.1, 1.3, and 22.2 s , respectively; the runtime for our $m w f$ algorithm was $39.1,368.4,449$ and 1317.6 s , respectively.

Table 4. Runtimes in seconds.

| Koblenz Datasets |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Dataset | $m h f$ |  |  | $m w f$ |  |  |
|  | $m h f-[10]$ | $m h f$-Ours | $m h f$-[10]/Ours | $m w f$-[10] | $m w f$-Ours | $m w f$-[10]/Ours |
| epin | 0.04 | $6.6 \times 10^{-3}$ | 6.106 | 0.04 | 0.02 | 2 |
| elec | $1.3 \times 10^{-2}$ | $3.6 \times 10^{-4}$ | 36 | 0.01 | 0.003 | 3.33 |
| fb | $1.8 \times 10^{-2}$ | $1.1 \times 10^{-3}$ | 15.9 | 0.017 | 0.03 | 0.56 |
| flickr | 5.57 | 0.45 | 12.34 | 5.65 | 1.37 | 4.12 |
| growth | 9.56 | 1.67 | 5.72 | 9.95 | 4.2 | 2.36 |
| youtube | 0.32 | $1.2 \times 10^{-2}$ | 25.8 | 0.33 | 0.18 | 1.83 |
| digg | $2.1 \times 10^{-3}$ | $9.1 \times 10^{-5}$ | 23.66 | 0.002 | 0.002 | 1 |
| slash | 0.01 | $1.3 \times 10^{-3}$ | 11.17 | 0.01 | 0.005 | 2 |
| conflict | $9 \times 10^{-5}$ | $3 \times 10^{-5}$ | 3 | 0.0009 | 0.02 | 0.045 |
| arxiv | $5 \times 10^{-2}$ | $9.3 \times 10^{-3}$ | 6.25 | 0.06 | 0.15 | 0.4 |
| enron | 0.18 | $2.8 \times 10^{-3}$ | 64.11 | 0.7 | 0.03 | 23.3 |
| delicious | 222 | 1.07 | 207.5 | 763 | 41 | 18.6 |
| Synthetic datasets $\mu_{I}=4, \mu_{D}=5, \mu_{T}=3$ |  |  |  |  |  |  |
| youtube | 120.8 | 0.17 | 679.5 | 186.8 | 71.3 | 2.6 |
| Synthetic runtimes $\mu_{I}=4, \mu_{D}=8, \mu_{T}=3$ |  |  |  |  |  |  |
| youtube | 132.7 | 0.29 | 451.3 | 275 | 95.5 | 2.8 |

mhf speedups over Bentert - koblenz datasets


Figure 13. mhf speedups on Koblenz datasets.


Figure 14. $m h f$ speedups on synthetic datasets.


Figure 15. mwf speedups on Koblenz datasets.


Figure 16. $m w f$ speedups on synthetic datasets.

## 8. Conclusions

We have shown that finding $m w f$ paths and walks as well as $m c f$ paths in intervaltemporal graphs is NP-hard. For the $m h f$ single-source all-destinations problem, a polynomial time algorithm was developed and for the $m w f$ single-source all-destinations problem, a pseudopolynomial time algorithm was developed. We show also that the $m h f$ and $m w f$ problems for interval graphs can be solved using the linear combination algorithm of Bentert et al. [10] for those interval graphs that can be modeled as contact-sequence graphs. While the algorithm of [10] is polynomial in the size of the contact-sequence graph, the modeling of an interval-temporal graph by a contact-sequence graph, when possible, may increase the graph size by an exponential amount and, so, this approach does not result in a polynomial time algorithm for interval-temporal graphs. In fact, even though all of our datasets could be modeled as contact-sequence graphs, we were unable to use the algorithm of [10] on one of our Koblenz datasets and three of our synthetic datasets as the code of [10] that transform the contact-sequence graph into the required input format failed due to insufficient memory. Our mhf paths algorithm outperformed the linear combination algorithm of [10] on all datasets on which the algorithm of [10] could be run. A speedup of up to 69.9 was obtained on the Koblenz datasets and up to 679 on synthetic datasets. For the $m w f$ single-source all-destinations problem, which is NP-hard, the linear combination algorithm of [10] outperformed our algorithm on three of the Koblenz data sets and tied on a fourth. On all remaining datasets, our algorithm outperformed that of [10]. The datasets on which the algorithm of [10] did well had the property that few vertices were reachable from the source vertex. This enabled the algorithm of [10] to terminate early. In these cases, the speedup obtained by our algorithm ranged from 0.045 to 0.56 . On the remaining Koblenz datasets, our $m w f$ algorithm obtained a speedup of up to 23.3. For both the synthetic datasets where it was possible to run the algorithm of [10], our algorithm was faster and delivered a speedup of up to 2.8 . On the four datasets that the transformation to the input format required by the algorithm of [10] failed due to lack of memory, our mhf algorithm ran in 1, 1.1, 1.3, and 22.2 s, respectively and the runtime for our $m w f$ algorithm was $39.1,368.4,449$ and 1317.6 s, respectively.


#### Abstract

Author Contributions: Conceptualization, A.J. and S.S.; Data curation, A.J. and S.S.; Formal analysis, A.J. and S.S.; Investigation, A.J. and S.S.; Methodology, A.J. and S.S.; Project administration, S.S.; Resources, A.J. and S.S.; Software, A.J.; Supervision, S S.; Validation, A.J. and S.S.; Visualization, A.J. and S.S.; Writing-original draft, A.J. and S.S.; Writing—review \& editing, A.J. and S.S. All authors have read and agreed to the published version of the manuscript.


Funding: This research received no external fund.
Institutional Review Board Statement: Not applicable.
Informed Consent Statement: Not applicable.
Data Availability Statement: The Koblenz datasets used for benchmarking are available in the KONECT graphs [18]. The social network graphs of youtube, flickr, livejournal are available at http://socialnetworks.mpi-sws.org/data-imc2007.html (accessed on 16 August 2022). This was shared by the authors of [17].

Acknowledgments: We acknowledge Wu et al. [8] and Bentert et al. [10] for sharing their implementation with us to benchmark against.

Conflicts of Interest: The authors declare that they have no competing interests.

## Abbreviations

The following abbreviations are used in this manuscript:

| $N A P P$ | No-wait acyclic path problem |
| :--- | :--- |
| $V$ | Number of vertices in a graph |
| $N$ | Used interchangeably with $V$ |
| $E_{S}$ | Number of edges in underlying static graph |
| $m w f$ | Min-wait foremost |
| $m h f$ | Mih-hop foremost |
| $m c f$ | Min-cost foremost |
| $M_{i t g}$ | Number of edges in interval-temporal graph (same as $E_{S}$ ) |
| $M_{c s g}$ | Number of contact-sequence edges |
| $\lambda$ | Travel duration on an edge at a given departure time |
| $\delta$ | Maximum number of departure intervals on an edge |

## References

1. Scheideler, C. Models and Techniques for Communication in Dynamic Networks. In Proceedings of the STACS19th Annual Symposium on Theoretical Aspects of Computer Science, Antibes, Juan les Pins, France, 14-16 March 2002; Volume 2285, pp. 27-49.
2. Stojmenović, I., Location Updates for Efficient Routing in Ad Hoc Networks. In Handbook of Wireless Networks and Mobile Computing; John Wiley \& Sons, Ltd.: Hoboken, NJ, USA, 2002; Chapter 21, pp. 451-471. [CrossRef]
3. Holme, P.; Saramäki, J. Temporal networks. Phys. Rep. 2012, 519, 97-125. [CrossRef]
4. Michail, O. An Introduction to Temporal Graphs: An Algorithmic Perspective. arXiv 2015, arXiv:1503.00278.
5. Santoro, N.; Quattrociocchi, W.; Flocchini, P.; Casteigts, A.; Amblard, F. Time-Varying Graphs and Social Network Analysis: Temporal Indicators and Metrics. arXiv 2011, arXiv:1102.0629.
6. Kuhn, F.; Oshman, R. Dynamic Networks: Models and Algorithms. SIGACT News 2011, 42, 82-96. [CrossRef]
7. Bhadra, S.; Ferreira, A. Computing multicast trees in dynamic networks and the complexity of connected components in evolving graphs. J. Internet Serv. Appl. 2012, 3, 269-275. [CrossRef]
8. Wu, H.; Cheng, J.; Ke, Y.; Huang, S.; Huang, Y.; Wu, H. Efficient Algorithms for Temporal Path Computation. IEEE Trans. Knowl. Data Eng. 2016, 28, 2927-2942. [CrossRef]
9. Bui-Xuan, B.M.; Ferreira, A.; Jarry, A. Evolving graphs and least cost journeys in dynamic networks. In Proceedings of the WiOpt'03: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks, Sophia Antipolis, France, 3-5 March 2003.
10. Bentert, M.; Himmel, A.S.; Nichterlein, A.; Niedermeier, R. Efficient computation of optimal temporal walks under waiting-time constraints. Appl. Netw. Sci. 2020, 5, 73. [CrossRef]
11. Guo, F.; Zhang, D.; Dong, Y.; Guo, Z. Urban link travel speed dataset from a megacity road network. Sci. Data 2019, 6, 61. [CrossRef] [PubMed]
12. Gheibi, S.; Banerjee, T.; Ranka, S.; Sahni, S. An Effective Data Structure for Contact Sequence Temporal Graphs. In Proceedings of the 2021 IEEE Symposium on Computers and Communications (ISCC), Athens, Greece, 5-8 September 2021; pp. 1-8. [CrossRef]
13. Jain, A.; Sahni, S. Min Hop and Foremost Paths in Interval Temporal Graphs. In Proceedings of the 2021 IEEE Symposium on Computers and Communications (ISCC), Athens, Greece, 5-8 September 2021; pp. 1-7.
14. Bhadra, S.; Ferreira, A. Complexity of Connected Components in Evolving Graphs and the Computation of Multicast Trees in Dynamic Networks. In Proceedings of the Ad-Hoc, Mobile, and Wireless Networks, Montreal, QC, Canada, 8-10 October 2003; Pierre, S., Barbeau, M., Kranakis, E., Eds.; Springer: Berlin/Heidelberg, Germany, 2003; pp. 259-270.
15. Casteigts, A.; Himmel, A.; Molter, H.; Zschoche, P. The Computational Complexity of Finding Temporal Paths under Waiting Time Constraints. arXiv 2019, arXiv:1909.06437.
16. Zschoche, P.; Fluschnik, T.; Molter, H.; Niedermeier, R. The Complexity of Finding Small Separators in Temporal Graphs. arXiv 2018, arXiv:1711.00963.
17. Mislove, A.; Marcon, M.; Gummadi, K.P.; Druschel, P.; Bhattacharjee, B. Measurement and Analysis of Online Social Networks. In Proceedings of the 5th ACM/Usenix Internet Measurement Conference (IMC’07), San Diego, CA, USA, 24-26 October 2007.
18. Kunegis, J. KONECT: The Koblenz Network Collection. In Proceedings of the 22nd International Conference on World Wide Web, WWW '13 Companion, Rio de Janeiro, Brazil, 13-17 May 2013; Association for Computing Machinery: New York, NY, USA, 2013; pp. 1343-1350. [CrossRef]
