



Restless reachability problems in temporal graphs

Suhas Thejaswi¹ · Juho Lauri² · Aristides Gionis³

Received: 9 September 2024 / Revised: 21 November 2024 / Accepted: 11 March 2025 /

Published online: 1 April 2025

© The Author(s) 2025

Abstract

We study a family of reachability problems under waiting-time restrictions in temporal and vertex-colored temporal graphs. Given a temporal graph and a set of source vertices, we find the set of vertices that are reachable from a source via a time-respecting path, where the difference in timestamps between consecutive edges is at most a resting time. Given a vertex-colored temporal graph and a multiset query of colors, we find the set of vertices reachable from a source via a time-respecting path such that the vertex colors of the path agree with the multiset query and the difference in timestamps between consecutive edges is at most a resting time. These kinds of problems have applications in understanding the spread of a disease in a network, tracing contacts in epidemic outbreaks, finding signaling pathways in the brain network, and recommending tours for tourists, among others. We present an algebraic algorithmic framework based on constrained multilinear sieving for solving the restless reachability problems we propose. In particular, parameterized by the length k of a path sought, we show that the proposed problems can be solved in $O(2^k km \Delta)$ time and $O(n\Delta)$ space, where n is the number of vertices, m the number of edges, and Δ the maximum resting time of an input temporal graph. The approach can be extended to extract paths and connected subgraphs in both static and temporal graphs, thus improving the work of Björklund et al. (in Proceedings of the European symposium on algorithms, 2014) and Thejaswi et al. (Big Data 8:335–362, 2020). In addition, we prove that our algorithms for the restless reachability problems in vertex-colored temporal graphs are optimal under plausible complexity-theoretic assumptions. Finally, with an open-source implementation, we demonstrate that our algorithm scales to large graphs with up to one billion temporal edges, despite the problems being NP-hard. Specifically, we present extensive experiments to evaluate our scalability claims both on synthetic and on real-world graphs. Our implementation is efficiently engineered and highly optimized. For instance, we can solve the restless reachability

The author contributed to this research during their employment at Aalto University, Finland.

✉ Suhas Thejaswi
thejaswi@mpi-sws.org

Juho Lauri
juho.lauri@gmail.com

Aristides Gionis
argioni@kth.se

¹ Max Planck Institute for Software Systems, Kaiserslautern, Germany

² Helsinki, Finland

³ Division of Theoretical Computer Science, KTH Royal Institute of Technology, Stockholm, Sweden

problem by restricting the path length to 9 in a real-world graph dataset with over 36 million directed edges in less than one hour on a commodity desktop with a 4-core Haswell CPU.

Keywords Algebraic fingerprinting · Multilinear sieving · Restless paths · Restless reachability · Temporal paths · Temporal reachability

1 Introduction

Graphs are used to model many real-world problems such as information propagation in social networks [3, 30], spreading of epidemics [2, 46, 48], protein interactions [1, 51], activity in brain networks [12, 20, 24], and design of nano-structures using DNA [4]. However, real-world problems often entail complex interactions whose semantics are not captured by usual simple graph models. As such, over the years, researchers have enriched graph models by introducing (i) node and edge attributes, giving rise to attributed graphs [47, 57] or (ii) edge timestamps, giving rise to temporal graphs [27, 43]. In particular, temporal graphs are used to model complex phenomena and network dynamics in a wide range of applications related to social networks, genealogical trees, transportation, and telecommunication networks.

As with any graph model, connectivity is a fundamental problem in temporal graphs, i.e., whether two vertices are connected by a time-respecting path or a temporal path—a path in which consecutive edges have non-decreasing timestamps [27, 43]. An extension to connectivity is reachability, where the goal is to find the (temporal) connectivity between all pairs of vertices [26]. Some variants of connectivity and reachability problems such as finding the path that minimizes the length or arrival time can be solved in polynomial time [16, 26, 60]. However, Casteigts et al. [13] showed that a variant of the connectivity problem with resting (or waiting) time restrictions—the time difference of consecutive edges is at most a resting time—known as the restless temporal path problem (or more simply a restless path if the context is clear) is NP-hard. In this work, we study a family of connectivity and reachability problems in temporal graphs with resting time restrictions. Specifically, we study the following problems.

Restless path problems. The restless path problem (RESTLESSPATH) asks if there exists a restless path connecting a source and a destination. Extending this, the short restless path problem (k - RESTLESSPATH) asks if there exists a restless path with a specified length, $k - 1$, connecting the source and the destination, and the short restless path motif problem (k - RESTLESSMOTIF) asks if there exists a restless path in a vertex-colored temporal graph, whose vertex colors match with the given multiset of colors of size k .

Restless reachability problems. The restless reachability problem (RESTLESSREACH) asks to find a set of vertices for which there exists a restless path connecting from a given source to the vertex. Extending this, the short restless reachability problem (k - RESTLESSREACH) asks to find a set of vertices for which there exists a restless path with a specified length, $k - 1$, connecting a given source to the vertex, and the short restless path motif reachability (k - RESTLESSMOTIFREACH) asks to find a set of vertices for which there exists a restless path whose vertex colors match with the given multiset of colors of size k , in a vertex-colored temporal graph.

Restless connectivity problems (i.e., restless walks and restless paths) have been extensively studied for estimating infections in an epidemic outbreak [13, 25]. Here, each vertex in the graph associates to a person and each temporal edge to a social interaction between two persons at a specific timestamp. The resting time of a vertex is viewed as the time until

which the virus/disease can propagate from that vertex—after exceeding the resting time, the virus/disease becomes inactive and stops propagating—in particular, the resting time of a vertex captures the recovery state of susceptible-infected-recovered (SIR) model of disease propagation [33]. Given the source of an infection, we want to find the set of people who may also be infected with the disease. Additionally, epidemiologists would be interested in a tool that allows them to evaluate different immunization strategies, such as computing efficiently the spread of the disease when a selected subset of the population has been immunized to the disease, e.g., vaccinated or quarantined. Beyond epidemiology, restless reachability has applications in finding signaling pathways in brain networks and tour recommendations. See Appendix A for further motivation.

All problems considered in this work are **NP**-hard, and as such, it is likely that they admit no polynomial-time algorithm. In such cases, it is typical to resort to heuristics or approximation algorithms, which still run in reasonable time but compromise the quality of the solution. In contrast, we consider exact algorithms for solving restless reachability problems in both temporal and vertex-colored temporal graphs. To also ensure fast runtime, our algorithms are exponential in the length of the path sought. More precisely, we show that when the path length is small enough our solution scales to massive graphs with up to one billion temporal edges.

It is worth noting that Casteigts et al. [13] studied the RESTLESSPATH problem from a complexity-theoretic point of view. For instance, they proved that the problem remains **NP**-hard on highly structured graphs such as complete graphs with exactly one edge removed. Despite several negative results, the authors pinpoint some parameters p of the problem for which the problem admits an algorithm running in time $f(p)n^{\mathcal{O}(1)}$, where $f(p)$ is some computable function depending solely on the parameter p . Namely, these parameters are the feedback edge number (FEN), the timed feedback vertex number (TFVN), and the length of the restless path. For FEN, the algorithm they describe runs in time $2^{\mathcal{O}(p)}n^{\mathcal{O}(1)}$ and for TFVN, the algorithm runs in time $6^q q! n^{\mathcal{O}(1)}$, where p is the FEN and q is the TFVN of the input n -vertex temporal graph.¹ In order for these algorithms to be scalable,² the corresponding parameter values should be small in practice on relevant instances. Unfortunately, as we show in Appendix D (see Table 7), there are real-world instances where this is not the case. Therefore, on such instances, it appears that the only known parameterized algorithm for the problem that has hope of being practical is one that pushes the unavoidable exponential dependency into the length of the restless path.³ Our key contributions in this paper are as follows:

- We present a generating function for generating restless walks, and a space-efficient algorithm based on constrained multilinear sieving [9] for solving k -RESTLESSMOTIF and k -RESTLESSMOTIFREACH in time $\mathcal{O}(2^k km \Delta)$ and $\mathcal{O}(n \Delta)$ space, where n is the number of vertices, m is the number of edges, $k - 1$ is the path length, and Δ is the maximum resting time. Furthermore, we show that our algorithm solves k -RESTLESSPATH and k -RESTLESSREACH in $\mathcal{O}(2^k km \Delta)$ time using $\mathcal{O}(n \Delta)$ space. Throughout this

¹ In the field of parameterized algorithms, when devising algorithms for **NP**-hard problems, it is usual to ignore the exact form of any polynomial factors in the run time and concentrate on the likely unavoidable exponential dependency on the parameter under consideration. That is, when the exact polynomial factor of the algorithm is not worked out (or omitted as non-essential), it is often represented as merely $|x|^{\mathcal{O}(1)}$ meaning $|x|^c$ for some $c \geq 1$, where x is the input and $|x|$ is its size.

² As the problem is **NP**-hard, the function $f(p)$ is necessarily exponential unless **P** = **NP**.

³ Casteigts et al. [13] also presented a $2^k n^{\mathcal{O}(1)}$ algorithm for RESTLESSPATH with respect to the parameter k , which is the length of the path. We compare our approach with theirs in Sect. 1.1.

paper, we call this algorithm a decision oracle as it returns a YES/NO answer with no explicit solution.

- Next, we develop the decision oracle into a fine-grained decision oracle capable of reporting the set of vertices that are reachable via a restless path from given source vertices along with the set of timestamps at which the vertices are reachable. Further, we exploit this to extract optimal solutions (i.e., solutions in which the maximum timestamp on the restless path is minimized) for k -RESTLESSMOTIF and k -RESTLESSPATH. Notably, our solution improves upon the earlier work of Thejaswi et al. [53, 54] by reducing the number of queries by a factor of $\log \tau$. Further, for extracting a solution, we reduce the number of queries from the work of Björklund et al. [8] from $\mathcal{O}(k \log n)$ to k . In total, our extraction algorithm runs in $\mathcal{O}(2^k km \Delta)$ time and $\mathcal{O}(n\tau)$ space, where τ is the maximum timestamp.
- As a consequence of our fine-grained decision oracle, our algorithm can answer more detailed queries. In particular, it can answer whether a given vertex u is reachable from source s at timestamp i with a restless path of length ℓ . Such a fine-grained oracle can be used to solve multiple variants of the restless path problem, such as finding a restless path that minimizes the path length, the arrival time, or the total resting time. Another key strength of our algorithm is its adaptability to multiple variants of the restless path problem. For instance, the algorithm is not limited to a single source, but can be extended to include multiple sources.
- We prove that our algorithms for k -RESTLESSMOTIF and k -RESTLESSMOTIFREACH are optimal under plausible complexity-theoretic assumptions. More precisely, we prove that there exists no $\mathcal{O}^*((2 - \epsilon)^k)$ -time algorithm⁴ for k -RESTLESSMOTIF or k -RESTLESSMOTIFREACH for any $\epsilon > 0$, assuming the so-called Set Cover Conjecture [18], a precise definition is given in Sect. 4.9.
- With an open-source implementation, we demonstrate that our algorithms scale to graphs with up to ten million edges on a commodity desktop equipped with an 4-core Haswell CPU. When scaled to a computing server with 2×12 -core Haswell CPU, the algorithm scales to graphs with up to one billion temporal edges [55]. As a case study, we model disease spreading in social gatherings as a k -RESTLESSREACH problem and perform experiments to study the propagation of the disease using the co-presence datasets from socio-patterns [23]. We also perform experiments to check the change in the disease-spreading pattern with the presence of people with immunity.

1.1 Related work

Our works builds upon the further related work on multilinear sieving and temporal reachability.

Multilinear sieving. Algebraic algorithms based on multilinear sieving for solving path problems in static graphs are due to Koutis and Williams [34, 35, 59]. Björklund et al. [5] improved the technique using narrow sieves to get a $\mathcal{O}^*(1.66^n)$ time algorithm for the Hamiltonian path problem. The sieving technique was extended by Björklund et al. [9] to find paths and subgraphs that agree with a multiset of colors. A practical implementation of the sieve was provided by Björklund et al. [10]. Furthermore, its parallelizability to vector-parallel architectures and scalability to large subgraph sizes was shown by Kaski et al. [32]. Note that

⁴ The notation \mathcal{O}^* hides the factors polynomially bounded in the input size.

these algorithms only handle static graphs. More recently, Thejaswi et al. [53, 54] extended the sieving technique to solve a family of pattern-detection problems in temporal graphs.

Reachability in temporal graphs. Path problems in temporal graphs are well-studied within different communities, including algorithms, data mining, and complex networks. The problem variants that seek to find a path that minimizes different objectives such as path length. In recent years, there has been an emphasis on more expressive temporal path problems due to their applicability in various fields [2, 12, 15, 22, 26, 48, 52, 58].

The problem of finding a temporal path under waiting-time restrictions was introduced by Dean [21], who studied a polynomial-time solvable variant of the problem. Himmel et al. [25] considered the restless walk problem where several visits to a vertex are allowed and presented polynomial-time algorithms. Thejaswi et al. [54][Section Path motif problem with delays] studied a variant of the temporal path problem that considers both transition time—the time required to transit from a vertex to its neighbor—and delay time—the minimum time that must be spent at each vertex before proceeding—and presented a $\mathcal{O}^*(2^k)$ algorithm. In their setting, the time difference of consecutive edges must be *at least* a vertex-specific delay time, in contrast to restless paths, where the time difference must be *at most* a vertex-specific waiting time.

More recently, Casteigts et al. [13] showed that RESTLESSPATH is **NP**-hard and also that k -RESTLESSPATH is **W**[1]-hard parameterized by either the feedback vertex number or the pathwidth of the input graph. The authors presented algorithms with running time $\mathcal{O}(2^k n^{\mathcal{O}(1)})$ and $\mathcal{O}(2^n n^{\mathcal{O}(1)})$ for k -RESTLESSPATH and RESTLESSPATH, respectively, where n is the number of vertices in the input temporal graph and k the length of the path.⁵ In comparison with Casteigts et al. [13], our approach can decide the existence of a solution in time $\mathcal{O}(2^k km \Delta)$ and $\mathcal{O}(2^n nm \Delta)$, where n is the number of vertices, m is number of temporal edges, and Δ is the maximum resting time. It should be noted that their algorithm only reports the existence of a restless path between a source and a destination with a YES/NO answer but does not return an explicit solution. In many practical scenarios, such as where further analysis of the path is required, this is insufficient. In a recent work building on our research, Zschoche [63] presented a randomized $\mathcal{O}(4^{k-d}(k-d)^2 m^3 \Delta)$ -time algorithm to find a shortest restless path (i.e., one minimizing the length of the path) between a source s and destination z , where d is the length of the shortest temporal path from s to z .

RESTLESSPATH and its variants studied in this work are self-reducible. By extending the approach of Björklund et al. [8], extracting an optimal k -restless path (i.e., one minimizing the maximum timestamp) can be done with $\mathcal{O}(k \log n \log \tau)$ queries to the decision oracle, followed by $\mathcal{O}(k!)$ steps to identify the path.⁶ However, using our fine-grained oracle we can extract a solution with $k - 1$ queries. As a result, an optimal restless path can be extracted in $\mathcal{O}(2^k km \Delta)$ and $\mathcal{O}(2^n nm \Delta)$ time for k -RESTLESSPATH and RESTLESSPATH, respectively. The RESTLESSREACH and k -RESTLESSREACH problems generalize RESTLESSPATH and k -RESTLESSPATH problems, respectively. With a decision oracle for RESTLESSPATH (or k -RESTLESSPATH), the set of reachable vertices from a source can be identified with $\mathcal{O}(n)$ queries, by repeatedly checking the existence of a restless path between source and each vertex. However, our method can answer reachability queries in a single computation, without

⁵ To further elaborate, Casteigts et al. [13] build upon the work of Koutis and Williams [35] using their result as black box. While this is sufficient to establish fixed-parameter tractability with respect to k , the exact polynomial factors are not fully detailed. Based on our calculations, the polynomial factor is $n^{\mathcal{O}(1)} \approx \mathcal{O}(kn + km \Delta)$ and the space complexity is $\mathcal{O}(kn\tau)$.

⁶ See Sect. 4.7 for a description of extracting a solution using self-reducibility.

having to iterate over each vertex. In addition, our method solves a general variant of restless path and restless reachability problems with color constraints on the vertices.

Organization of the paper. The rest of this paper is organized as follows: We introduce the necessary notation in Sect. 2, and we formally define the reachability problems for temporal graphs and vertex-colored temporal graphs in Sect. 3. Afterward, our algorithmic solution is presented in Sect. 4. Our empirical evaluation on a large collection of synthetic and real-world graphs is discussed in Sect. 5. Finally, Sect. 6 offers a short conclusion, limitations, and directions for future work.

2 Terminology

In this section, we introduce our notation and basic terminology. For a positive integer k , we write $[k] = \{1, \dots, k\}$. A list of symbols used in this paper is available in Appendix B Table 6.

Static graphs. A static undirected graph G is a tuple (V, E) where V is a set of vertices and E is a set of unordered pairs of vertices called edges. A static directed graph G is a tuple (V, E) where V is a set of vertices and E is a set of ordered pair of vertices called edges. A static walk between any two vertices is an alternating sequence of vertices and edges $u_1 e_1 u_2 \dots e_k u_k$ such that there exists an edge $e_i = (u_i, u_{i+1}) \in E$ for each $i \in [k - 1]$. We call the vertices u_1 and u_k the start and end vertices of the walk, respectively. We refer to walk W as (s, d) -walk for $u_1 = s$ and $u_k = d$. The length of a static walk is the number of edges in the walk. A static path is a static walk with no repetition of vertices.

Undirected temporal graphs. An undirected temporal graph G is a tuple (V, E) , where V is a set of vertices and E is a set of undirected temporal edges. An undirected temporal edge is a tuple (u, v, j) where $u, v \in V$ and $j \in [\tau]$ is a timestamp, where τ is the maximum timestamp in G . Note that, by definition, two undirected edges (u, v, j) and (v, u, j) are equivalent. A vertex u is adjacent to vertex v and vice versa at timestamp i in an undirected graph G if there exists an undirected edge $(u, v, i) \in E$. A temporal walk W is an alternating sequence of vertices and temporal edges $u_1 e_1 u_2 e_2 \dots e_{k-1} u_k$ such that $e_i \in E$ for all $i \in [k - 1]$ and for any two edges $e_i = (u_i, u_{i+1}, j_i)$, $e_{i+1} = (u_{i+1}, u_{i+2}, j_{i+1})$ in W , it is $j_i \leq j_{i+1}$. We say the walk W reaches vertex u_k at time j_{k-1} . The vertices u_1 and u_k are called source and destination vertices of W , respectively. The vertices $\{u_2, \dots, u_{k-1}\}$ are called in-vertices (or equivalently, internal vertices) of W . We refer to the temporal walk W as (s, d) -temporal walk with source $s = v_1$ and destination $d = v_k$. The vertex set and edge set of walk W is denoted as $V(W)$ and $E(W)$, respectively. The length of a temporal walk is the number of edges in the temporal walk. A temporal path is a temporal walk with no repetition of vertices.

Directed temporal graphs. A directed temporal graph G is a tuple (V, E) , where V is a set of vertices and $E \subseteq V \times V \times [\tau]$ is a set of directed temporal edges at discrete time steps, where τ is the maximum timestamp. A directed temporal edge is a tuple (u, v, j) where $u, v \in V$ and $j \in [\tau]$ is a timestamp. A directed edge (u, v, j) is referred as an outgoing or departing edge for u and an incoming or arriving edge for v at time j . A vertex u is an in-neighbor to v at time j if $(u, v, j) \in E$, similarly a vertex v is an out-neighbor to u at time j if $(u, v, j) \in E$. The set of in-neighbors to v at time j is denoted by $N_j(v) = \{u \mid (u, v, j) \in E\}$, and the set of out-neighbors to u at time j is denoted by $N'_j(u) = \{v \mid (u, v, j) \in E\}$.

Unless stated otherwise, by a temporal graph we always mean a directed temporal graph. Similarly, by a temporal walk we always mean a directed temporal walk, and by a temporal

path we always mean a directed temporal path. To distinguish a temporal graph from a static graph, we sometimes explicitly call a graph non-temporal or static.

Restless walk. A restless temporal walk, or simply restless walk, is a temporal walk such that for any two consecutive edges $e_i = (u_i, u_{i+1}, j_i)$ and $e_{i+1} = (u_{i+1}, u_{i+2}, j_{i+1})$, it holds that $j_{i+1} - j_i \leq \delta(u_{i+1})$, where the function $\delta : V \rightarrow [\Delta]$ defines a vertex-dependent waiting time, with $\Delta \in \mathbb{N}_+$ being the maximum waiting time. In other words, in a restless temporal walk, it is not allowed to wait more than a vertex-dependent amount of time in each vertex. A restless path is a restless walk with no repetition of vertices.

Coloring. Let C to be a set of colors. A vertex-colored temporal graph $G = (V, E)$ is a temporal graph with a vertex-coloring function $c : V \rightarrow C$, which maps each vertex $u \in V$ to a subset of colors in C . Let M be a multiset of colors and W be a walk. The walk W is properly colored (or is said to agree with M) if there exists a bijection $f : V(W) \rightarrow M$ such that $\{f(u)\} \cap \{c(u)\} \neq \emptyset$ for all $u \in V(W)$.

Polynomials. We assume that our temporal graph contains n vertices and m temporal edges, and for simplicity we write $V = \{u_1, \dots, u_n\}$ and $E = \{e_1, \dots, e_m\}$. We introduce a variable x_{u_i} for each $u_i \in V$ and a variable y_{e_j} for each $e_j \in E$. Let \mathcal{P} be a multivariate polynomial such that each monomial $\mathcal{P} \in \mathcal{P}$ is of the form $x_{u_1}^{f_1} \dots x_{u_n}^{f_n} y_{e_1}^{g_1} \dots y_{e_m}^{g_m}$. A monomial \mathcal{P} is multilinear if $f_i, g_j \in \{0, 1\}$ for all $i \in [n]$ and $j \in [m]$. The degree (or size) of P is the sum of the degrees of all the variables in \mathcal{P} . Let C be a set of colors and $c : V \rightarrow C$ be a vertex-coloring function. Let M be a multiset of colors. For each $s \in C$, let $\mu(s)$ denote the number of occurrences of color s in M , noting that $\mu(s) = 0$ if $s \notin M$. We say that a monomial $x_{u_1}^{f_1} \dots x_{u_n}^{f_n} y_{e_1}^{g_1} \dots y_{e_m}^{g_m}$ is properly colored if for each $s \in C$ it holds that $\mu(s) = \sum_{i \in c^{-1}(s)} d_i$, in other words, the number of occurrences of color s is equal to the total degree of x -variables representing the vertices with color s .

3 Restless path and restless reachability problems

In this section, we define the problems that we consider in this paper. While we define the problems on temporal directed graphs, we note that our algorithmic approach can be extended to undirected graphs by replacing each undirected edge with two directed edges in opposite directions. In such case, the asymptotic complexity of the algorithm remains the same.

Restless path problem (RESTLESSPATH). Given a temporal graph $G = (V, E)$, a function $\delta : V \rightarrow \mathbb{N}_+$, a source $s \in V$, and a destination $d \in V$, the problem asks if there exists a restless path from s to d in G .

Short restless path problem (k - RESTLESSPATH). Given a temporal graph $G = (V, E)$, a function $\delta : V \rightarrow \mathbb{N}_+$, a source $s \in V$, a destination $d \in V$, and an integer $k \leq n$, the problem asks if there exists a restless path of length $k - 1$ from s to d in G .

Short restless path motif problem (k- RESTLESSMOTIF). Given a temporal graph $G = (V, E)$ with a coloring function $c : V \rightarrow C$ where C is a set of colors, a function $\delta : V \rightarrow \mathbb{N}_+$, a source $s \in V$, a destination $d \in V$, and a multiset M of colors, $|M| = k$, the problem asks if there exists a restless path from s to d in G such that the vertex colors of the path agree with M .

An illustration of restless path problems is available in Fig. 1. All these three problems are NP-hard (the hardness of the first two can be found in the paper of Casteigts et al. [13]). For

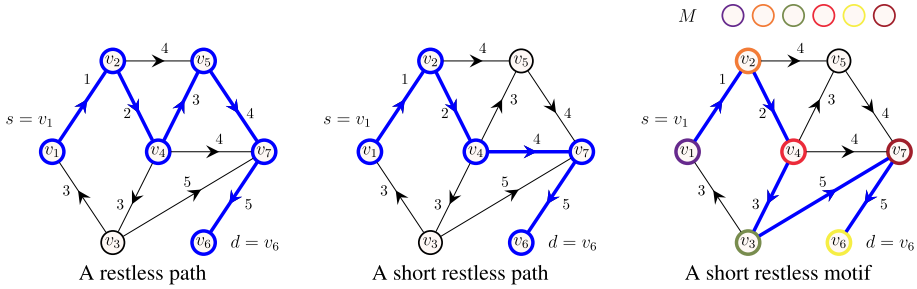


Fig. 1 Illustration of restless path problems. A temporal graph on vertices $V = \{v_1, \dots, v_7\}$ with source $s = v_1$, destination $d = v_6$ and resting time of vertices $\delta(v_1) = \dots = \delta(v_7) = 2$. Arrows represent the direction of edges and the integer value on each edge corresponds to its timestamp. On the left, an example of a restless path from v_1 to v_6 of length 5. On the center, an example of a restless path from v_1 to v_6 when the length of the path is restricted to 4 i.e., $k = 5$. On the right, an example of a short restless (path) motif from v_1 to v_6 such that the vertex colors of the path agree with the multiset of colors in M . Restless paths are highlighted in bold (blue)

the last claim, observe that k - RESTLESSPATH is a special case of k - RESTLESSMOTIF where all the vertices in the graph are colored with a single color and the query multiset is $M = \{1^k\}$. Since k - RESTLESSPATH is NP-hard, it follows that k - RESTLESSMOTIF is NP-hard, as well.

Restless reachability problem (RESTLESSREACH). Given a temporal graph $G = (V, E)$, a function $\delta : V \rightarrow \mathbb{N}_+$, and a source vertex $s \in V$, the problem asks to find the set of vertices $D \subseteq V$ such that for each $d \in D$ there exists a restless path from s to d in G . Clearly, the problem generalizes RESTLESSPATH and is thus computationally hard.

Short restless reachability problem (k - RESTLESSREACH). Given a temporal graph $G = (V, E)$, a function $\delta : V \rightarrow \mathbb{N}_+$, a source vertex $s \in V$, and an integer $k \leq n$, the problem asks to find the set of vertices $D \subseteq V$ such that for each $d \in D$ there exists a restless path of length $k - 1$ from s to d in G . Given that k - RESTLESSPATH is hard, k - RESTLESSREACH remains hard, as well.

Short restless motif reachability problem (k- RESTLESSMOTIFREACH). Given a temporal graph $G = (V, E)$ with coloring function $c : V \rightarrow C$ where C is a set of colors, a function $\delta : V \rightarrow \mathbb{N}_+$, and a multiset M of colors such that $|M| = k$, the problem asks to find the set of vertices $D \subseteq V$ such that for each $d \in D$ there exists a restless path from s and d in G such that the vertex colors of the path agree with multiset M . Again, a routine observation shows that this problem is computationally hard.

4 Algorithm

Our approach makes use of the polynomial encoding of temporal walks introduced by Thejaswi et al. [53], building on the earlier work of Björklund et al. [9] for static graphs. Because algebraic fingerprinting techniques are not well known, we first provide some intuition behind our methods before diving into details.

4.1 A birds-eye view of our approach

Algebraic fingerprinting methods have been successfully applied to pattern-detection problems on graphs. On a high level, the approach works by encoding a problem instance G as a multivariate polynomial where the variables in the polynomial correspond to entities of G such as a vertex or an edge. The key idea is to design a polynomial encoding, which evaluates to a nonzero term if and only if the desired pattern is present in G . We can then evaluate the polynomial by substituting random values to its variables. Now, if one of the substitution evaluates to a nonzero term, it implies that the desired pattern is present in G . Since the substitutions are random, the approach can give false negatives. However, the resulting algorithms typically have one-sided error meaning false positives are not possible. Moreover, as is typical, the error probability can be made arbitrarily small (e.g., as unlikely as hardware failure) by repeating the substitutions.

In our case, we need to decide the existence of a restless path in a temporal graph G . Following the approach presented above, we encode all the restless walks in G using a multivariate polynomial where each monomial corresponds to a restless walk and variables in a monomial correspond to vertices and/or edges. Specifically, we design an encoding where a monomial is multilinear (i.e., no variables in the monomial are repeated) if and only if the corresponding restless walk it encodes is a restless path. As such, the problem of deciding the existence of a restless path is equivalent to the problem of deciding the existence of a multilinear monomial in the generated polynomial. The existence of a multilinear monomial in a polynomial, in turn, can be determined using polynomial identity testing, in particular, by evaluating random substitutions for the variables; if one of the substitution evaluates to a nonzero term then there exists at least one multilinear monomial in the polynomial. Most importantly, *note that an explicit representation of the polynomial encoding can be exponentially large. However, in our approach we do not need to represent the polynomial explicitly, but we only need to evaluate random substitutions efficiently.*

For an interested reader, we present a detailed discussion on polynomial encoding of walks in static graphs and temporal graphs in Appendix C.

4.2 An overview of the algorithm

To obtain our algorithm, we proceed by taking the following steps. First, we present a dynamic-programming algorithm to encode restless walks as a polynomial and from this polynomial, showing that detecting a multilinear monomial of degree $2\ell - 1$ is equivalent to detecting a restless path of length $\ell - 1$, and vice versa. Then, we extend the approach to detecting restless paths with additional color constraints on the vertices. Finally, we use this algorithm to solve the problems described in Sect. 3.

In this work, we use finite fields for evaluating polynomials, in particular we make use of Galois fields (GF) of characteristic 2. An element of $\text{GF}(2^b)$ can be represented as a bit vector of length b . We can perform field operations such as addition and multiplication on these bit vectors in time $\mathcal{O}(b \log b)$ [41]. The polynomial $\chi_{u,\ell,i}(\vec{x}, \vec{y})$ on variables $\vec{x} = \{x_v : v \in V\}$, $\vec{y} = \{y_{uv,\ell,j} : (u, v, j) \in E, \ell \in [k]\}$ in Eq. (1) can be evaluated using a random assignment $\tilde{x} = \{\tilde{x}_v \in \text{GF}(2^b)\} : x_v \in \vec{x}$, $\tilde{y} = \{\tilde{y}_{uv,\ell,j} \in \text{GF}(2^b) : y_{uv,\ell,j} \in \vec{y}\}$. Now we can build an arithmetic circuit that represents an algorithm which evaluates the polynomial $\chi_{u,\ell,i}(\vec{x}, \vec{y})$ for input (\tilde{x}, \tilde{y}) in time linear in the number of gates in the circuit. Most importantly, observe that the expanded expression of the polynomial $\chi_{u,\ell,i}(\vec{x}, \vec{y})$ can be exponentially large; however, the arithmetic circuit evaluating $\chi_{u,\ell,i}(\vec{x}, \vec{y})$ can be reduced to size polynomial in number

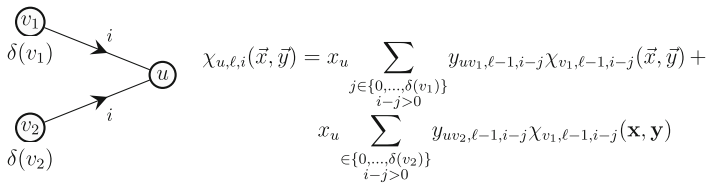


Fig. 2 An illustration of generating restless walks

of gates by applying the associativity of addition and distributivity of multiplication over addition.

For detecting the existence of a multilinear monomial in the generated polynomial, the algorithm works by randomly substituting values $(\vec{x}, \vec{y}) \in \text{GF}(2^b)$, for a suitable b , to variables in $\vec{x} = \{x_v : v \in V\}$ and $\vec{y} = \{y_{uv, \ell, j} : (u, v, j) \in E^\tau, \ell \in [k]\}$. Specifically, the parameter b is the number of bits used to generate the field variables. Because multiplication between two field variables is defined as an XOR operation, multiplication between two variables with the same value results in a zero term. That is, the monomials corresponding to walks that are not paths evaluate to a zero term, while the monomials corresponding to paths evaluate to a nonzero term. However, it is possible that an evaluation results in a zero term even though there is a multilinear monomial term corresponding to a path.⁷ To eliminate the effect of such false negatives, we repeat the evaluation of the polynomial with 2^k random substitutions. The probability of a false negative then becomes $2^{-b}(2k - 1)$, where $k - 1$ is the length of the temporal walk. In practice, the false-negative probability depends on the quality of the random number generator. It is important to note, however, that the false-positive probability is zero. For full technical details, we refer the interested reader to Björklund et al. [9] and Cygan et al. [19, Chapter 10].

4.3 Generating restless walks

In this section, we extend the dynamic-programming recursion to generate a polynomial encoding of restless walks of length $\ell - 1$ using encoding of restless walks of length $\ell - 2$.

An example is illustrated in Fig. 2, in which we depict a vertex u with incoming neighbors $N_i(u) = \{v_1, v_2\}$. From the definition of a restless walk, it is clear that we can continue the walk from vertices v_1, v_2 to vertex u at time i only if we had reached v_1 or v_2 no earlier than time $i - \delta(v_1)$ and $i - \delta(v_2)$, respectively. Let $\chi_{u, \ell, i}(\vec{x}, \vec{y})$ denote the encoding of all restless walks of length $\ell - 1$ ending at vertex u at time i .

We generate restless walks of length $\ell - 1$ that end at vertex u at time i using restless walks of length $\ell - 2$ by having

$$\chi_{u, \ell, i}(\vec{x}, \vec{y}) = x_u \sum_{\substack{j \in \{0, \dots, \delta(v_1)\} \\ i-j > 0}} y_{uv_1, \ell-1, i-j} \chi_{v_1, \ell-1, i-j}(\vec{x}, \vec{y}) + x_u \sum_{\substack{j \in \{0, \dots, \delta(v_2)\} \\ i-j > 0}} y_{uv_2, \ell-1, i-j} \chi_{v_2, \ell-1, i-j}(\vec{x}, \vec{y}).$$

⁷ A polynomial with degree $2k + 1$ has at most $2k + 1$ roots, so there is a chance that the chosen values of \vec{x} and \vec{y} are roots resulting in the polynomial to evaluate to zero even when the desired structure is present in the graph and the false-negative probability is bounded by Schwartz–Zippel lemma [50, 62].

Generalizing from the previous example, the dynamic-programming recursion is written as

$$\begin{aligned} \chi_{u,1,i}(\vec{x}, \vec{y}) &= x_u, \text{ for each } u \in V \text{ and } i \in [\tau], \text{ and} \\ \chi_{u,\ell,i}(\vec{x}, \vec{y}) &= x_u \sum_{v \in N_i(u)} \sum_{\substack{j \in \{0, \dots, \delta(v)\} \\ i-j > 0}} y_{uv, \ell-1, i-j} \chi_{v, \ell-1, i-j}(\vec{x}, \vec{y}), \end{aligned} \tag{1}$$

for each $u \in V, \ell \in \{2, \dots, k\}$ and $i \in [\tau]$.

The following result is fundamental to our approach.

Lemma 1 *The polynomial encoding $\chi_{u,\ell,i}(\vec{x}, \vec{y})$ presented in Eq. (1) contains a multilinear monomial of degree $2\ell - 1$ if and only if there exists a restless path of length $\ell - 1$ ending at vertex u reaching at time i .*

Proof We prove the claim by induction. For $l = 1, \chi_{u,1,i}(\vec{x}, \vec{y}) = x_u$ for every $u \in V$ and $i \in [\tau]$, so the base case holds trivially. Assume that for $\ell - 1, \chi_{u,\ell-1,i}(\vec{x}, \vec{y})$ contains a multilinear monomial of degree $2\ell - 3$ if and only if there exists a restless path of length $\ell - 2$ ending at vertex $u \in V$ at time $i \in [\tau]$.

For $\chi_{u,\ell,i}(\vec{x}, \vec{y})$, from Eq. (1), the y -variables in any monomial are not repeated, as they have unique ℓ and i subscript. By construction, we include the restless walks from neighbors $v \in N_i(u)$ to construct $\chi_{u,\ell,i}(\vec{x}, \vec{y})$, if v was reached no earlier than $i - \delta(v)$. Thus, all walks included in the polynomial encoding $\chi_{u,\ell,i}(\vec{x}, \vec{y})$ are restless. Moreover, by construction, all walks encoded in $\chi_{u,\ell,i}(\vec{x}, \vec{y})$ end at vertex u and have length $\ell - 1$, since all walks in $\chi_{v,\ell-1,i}(\vec{x}, \vec{y})$ have length $\ell - 2$. Suppose $\chi_{u,\ell,i}(\vec{x}, \vec{y})$ contains a multilinear monomial with degree $2\ell - 1$. Since all y -variables are unique, there must be ℓ unique x -variables, corresponding to restless path of length $\ell - 1$. Because, a monomial is multilinear if and only if it corresponds to a path, as established in [53, Lemma 2].

Conversely, suppose there exists a restless path of length $\ell - 1$ ending at vertex u at the latest time i , which is encoded as $x_1 \dots x_v y_{uv, \ell-1, i} x_u$, but $\chi_{u,\ell,i}(\vec{x}, \vec{y})$ does not contain this multilinear monomial of degree $2\ell - 1$. From the construction in Eq. 1, this implies that $\chi_{v,\ell-1,j}(\vec{x}, \vec{y})$ for any $j \leq i - \delta(v)$ must not contain the multilinear monomial $x_1 \dots x_v$ with degree $2\ell - 3$, which corresponds to a restless path of length $\ell - 2$ ending vertex v and reaching at time j . This contradicts the premise that $\chi_{v,\ell-1,j}(\vec{x}, \vec{y})$ contains a multilinear monomial of degree $2\ell - 3$ if and only if a restless path of length $\ell - 2$ exists, completing the proof. \square

The next part of the problem is to detect a multilinear monomial in the polynomial in Eq. (1) representing the restless walks. There is already established theory related to this problem [9, 34, 35]. In particular, it is known that by substituting 2^ℓ random values for the variables in x and y , and evaluating them, if one of the evaluation results in a nonzero term, it implies that there exists at least one multilinear monomial of degree $2\ell - 1$.

For each $u \in V$ and $j \in [\ell]$, we introduce a new variable $z_{u,j}$. The vector of all variables of $z_{u,j}$ is denoted as \mathbf{z} and the vector of all variables of y is denoted as \mathbf{y} . We write $z_u^L = \sum_{j \in L} z_{u,j}$, for $u \in V, L \subseteq [\ell]$ and $\mathbf{z}^L = \{z_u^L : u \in V\}$ for $L \subseteq [\ell]$. The values of the variables $z_{u,j}$ are assigned uniformly at random from $\text{GF}(2^b)$. For simplicity, we write $V = \{u_1, \dots, u_n\}$.

Lemma 2 (Multilinear sieving [7]) *The polynomial*

$$\zeta_{u,\ell,i}(\vec{z}, \vec{y}) = \sum_{L \subseteq [\ell]} \chi_{u,\ell,i}(z_{u_1}^L, \dots, z_{u_n}^L, \vec{y}) \tag{2}$$

is not identically zero if and only if $\chi_{u,\ell,i}(\vec{x}, \vec{y})$ contains a multilinear monomial of degree $2\ell - 1$.

Lemma 3 *Evaluating the polynomial in Eq. (2) can be done in time $\mathcal{O}(2^\ell \ell m \Delta)$ and space $\mathcal{O}(n \Delta)$.*

Proof Recall that $\Delta = \max_{v \in V} \delta(v)$, m_i is the number of edges at $i \in [\tau]$, and n is the number of vertices. Computing $\chi_{u,j,i}(\vec{x}, \vec{y})$ for all $u \in V$ requires $(\Delta + 1)m_i$ multiplications and additions. We repeat this for all $i \in [\tau]$ and $j \in [\ell]$, which requires $\mathcal{O}(\ell m \Delta)$ multiplications and additions. Finally, we evaluate the polynomial $\zeta_{u,\ell,i}(\vec{z}, \vec{y})$ for all $u \in V$ using 2^ℓ random substitution of variables in $\mathbf{z}^L = \{z_{u_1}^L, \dots, z_{u_n}^L\}$, for each $L \subseteq [\ell]$, which takes time $\mathcal{O}(2^\ell \ell m \Delta)$. So the run time is $\mathcal{O}(2^\ell \ell m \Delta)$. \square

In the next section, we introduce color constraints for the vertices in the restless path. More precisely, given a vertex-colored temporal graph $G = (V, E)$ with a coloring function $c : V \rightarrow C$, where C is a set of colors, and a multiset M of colors, we consider the problem of deciding the existence of a restless path such that the vertex colors of the path agree with the multiset of colors in M . This generalization of the restless path problem with color constraints will be used to solve restless reachability problems in the later sections.

4.4 Introducing vertex-color constraints

Given a multiset of colors M , we extend the multilinear sieving technique to detect the existence of a multilinear monomial, such that colors corresponding to the x -variables agree with the colors in multiset M . Recall our earlier definition where $\mu(s)$ denotes the number of occurrences of color s in a multiset M . Furthermore, for each $s \in C$, let S_s denote the set of $\mu(s)$ shades of the color s , with $S_s \cap S_{s'} = \emptyset$ for all $s \neq s'$. In other words, for each color $s \in C$ we create $\mu(s)$ shades, so that any two distinct colors have different shades. As an example, for the color multiset $M = \{1, 1, 2\}$ we have $S_1 = \{1_1, 1_2\}$ and $S_2 = \{2_1\}$, so that $S_1 \cap S_2 = \emptyset$.

For each $u \in V$ and $d \in S_{c(u)}$, we introduce a new variable $\gamma_{u,d}$. For each $d \in \cup_{s \in C} S_s$ and each label $j \in [\ell]$, we introduce a new variable $\omega_{j,d}$. The values of variables $\gamma_{u,d}$ and $\omega_{d,j}$ are drawn uniformly at random from the Galois field $\text{GF}(2^b)$. We write

$$z_{u,j} = \sum_{d \in S_{c(u)}} \gamma_{u,d} \omega_{d,j}, \text{ and } z_u^L = \sum_{j \in L} z_{u,j}, \tag{3}$$

for $L \subseteq [\ell]$, $u \in V$. The following lemma extends Lemma 2 in the case of vertex-color constraints.

Lemma 4 (Constrained multilinear sieving [9]) *The polynomial $\chi_{u,\ell,i}(\vec{x}, \vec{y})$ contains a multilinear monomial of degree $2\ell - 1$, and it is properly colored if and only if the polynomial*

$$\zeta_{u,\ell,i}(\vec{z}, \vec{w}, \vec{y}) = \sum_{L \subseteq [l]} \chi_{u,\ell,i}(z_{u_1}^L, \dots, z_{u_n}^L, \vec{y}) \tag{4}$$

is not identically zero.

From Lemma 4, we can determine the existence of a multilinear monomial in $\zeta_{u,\ell,i}(\vec{z}, \vec{w}, \vec{y})$, by making 2^ℓ random substitutions of the new variables \mathbf{z} in Eq. (4). As detailed in Björklund et al. [9], these substitutions can be random for a low-degree polynomial that is not identically zero with only few roots. If one evaluates the polynomial at a

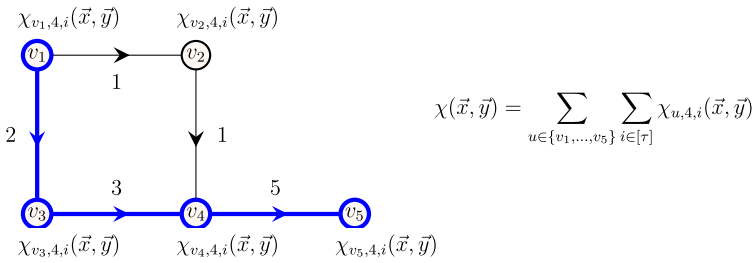


Fig. 3 An example to illustrate the fine-grained evaluation scheme. The resting time of vertices is $\delta(v_1) = \dots = \delta(v_5) = 2$. A restless path of length 3 (i.e., $\ell = 4$) from vertex v_1 to vertex v_5 is highlighted in bold

random point, one is likely to witness that it is not identically zero [50, 62]. This gives rise to a randomized algorithm with a false-negative probability of $2^{-b}(2\ell - 1)$, where the arithmetic is over the Galois field $GF(2^b)$. Here, again, b is the number of bits used to generate the random values. We make use of this result for designing our algorithms.

Lemma 5 *Evaluating the polynomial in Eq. (4) can be done in time $\mathcal{O}(2^\ell \ell m \Delta)$ and space $\mathcal{O}(n \Delta)$.*

The proof of Lemma 5 is similar to the one of Lemma 3. From Lemma 4, we obtain an algorithm to detect the existence of a restless path ending at vertex $u \in V$ at time i and the vertex colors of the restless path agree with the colors in multiset M .

4.5 A fine-grained decision oracle

In this section, we present a *fine-grained* evaluation scheme to evaluate the polynomial in Eq. (4).

Intuition. Consider the graph illustrated in Fig. 3, with vertex set $V = \{v_1, \dots, v_5\}$ with resting time $\delta(v_1) = \dots = \delta(v_5) = 2$, i.e., we can wait at most 2 time steps at any vertex. In order to decide whether there exists a restless path of length 3 (i.e., $\ell = 4$) in the graph, we need to evaluate the polynomial $\chi(\vec{x}, \vec{y}) = \sum_{u \in \{v_1, \dots, v_5\}} \sum_{i \in [\tau]} \chi_{u,4,i}(\vec{x}, \vec{y})$. However, to decide if there exists a restless path of length 3 ending at vertex v_5 , it is sufficient to evaluate the polynomial $\sum_{i \in [\tau]} \chi_{v_5,4,i}(\vec{x}, \vec{y})$. Furthermore, to decide whether there exists a restless path of length 3 ending at vertex v_5 at time 5, we can restrict the evaluation to the polynomial $\chi_{v_5,4,5}(\vec{x}, \vec{y})$. Similarly, it suffices to evaluate the polynomial $\chi_{u,\ell,i}(\vec{x}, \vec{y})$ to determine whether there exists a path of length $\ell - 1$ ending at vertex u at time i .

Fine-grained decision oracle. Let us generalize the fine-grained evaluation scheme to any temporal graph using the intuition presented. Instead of evaluating a single polynomial $\chi(\mathbf{x}, \mathbf{y}) = \sum_{u \in V} \sum_{\ell \in [k]} \sum_{i \in [\tau]} \chi_{u,\ell,i}(\vec{x}, \vec{y})$, we work with a set of $n\tau$ polynomials $\{\chi_{u,\ell,i}(\vec{x}, \vec{y}) : u \in V, i \in [\tau]\}$ and evaluate each $\chi_{u,\ell,i}(\vec{x}, \vec{y})$ independently. Observe carefully that our generating function in Eq. (1) generates a polynomial encoding of all restless walks $\chi_{u,\ell,i}(\vec{x}, \vec{y})$ for each $u \in V$ and $i \in [\tau]$ independently for a fixed ℓ . If the corresponding evaluation polynomial $\zeta_{u,\ell,i}(\vec{z}, \vec{w}, \vec{y})$ evaluates to a nonzero term, it implies that there exists a restless path of length $\ell - 1$ ending at vertex u at time i and the vertices in the path agree with the multiset of colors in M . Using this fine-grained evaluation scheme, we obtain a set of timestamps $R_u = \{i : \zeta_{u,\ell,i}(\vec{z}, \vec{w}, \vec{y}) \neq 0 \text{ for all } i \in [\tau]\}$ of restless paths ending

at vertex $u \in V$ and satisfying the color constraints in M . The pseudocode is presented in Algorithm 1.

The term *fine-grained oracle* is used to differentiate it from a decision oracle, which only reports the existence of a restless path with a YES/NO answer. Our fine-grained decision oracle reports more insightful details; for instance, it can answer if there exists a restless path of length $\ell - 1$ ending at vertex $u \in V$ at time $i \in [\tau]$ satisfying the color constraints specified in M with a YES/NO answer. A single run of the fine-grained oracle is sufficient to obtain the set of vertices $D \subseteq V$ such that there exists a restless path of length $\ell - 1$ ending at each $u \in D$.

This re-design of the polynomial evaluation scheme improves the run time for solving the temporal path problems described in previous work [53, 54], where the authors search for a temporal path that contains colors specified in the query. That is, by replacing their decision oracle with our fine-grained decision oracle, we reduce the number of queries by a factor of $\log \tau$. This, in turn, reduces the total runtime of their solution for detecting an optimal solution from $\mathcal{O}(2^\ell \ell(n\tau + m) \log \tau)$ to $\mathcal{O}(2^\ell \ell(n\tau + m))$. Even though the theoretical improvement is modest, it is important to note that for large values of ℓ a single run of the decision oracle can take hours to complete, so the practical improvement is significant (for precise results, see Sect. 5 and Table 4, as well as previous work [32]).

Algorithm 1: FINEGRAINEDORACLE($G = (V, E), c, k, M$)

```

Input:  $G = (V, E)$ , input graph
          $c$ , coloring function
          $k$ , length of path
          $M$ , multiset of colors
Output:  $R$ , minimum reachability time
          $D$ , set of reachable vertices
1 for  $L \subseteq [k]$  do
2    $\mathbf{z}^L \leftarrow \text{GET-Z-ASSIGNMENT}(c, M, L)$  // Eq. (3)
3   for  $u \in V, i \in \tau$  do
4      $\zeta_{u,1,i}(\vec{z}, \vec{w}, \vec{y}) \leftarrow z_u^L$ 
5 for  $\ell \in \{2, \dots, k\}$  do
6   for  $u \in V, i \in [\tau]$  do
7      $\zeta_{u,\ell,i}(\vec{z}, \vec{w}, \vec{y}) \leftarrow 0$ 
8     for  $L \subseteq [k]$  do
9        $\mathbf{z}^L \leftarrow \text{GET-Z-ASSIGNMENT}(c, M, L)$  // Eq. (3)
10       $\zeta_{u,\ell,i}(\vec{z}, \vec{w}, \vec{y}) \leftarrow \zeta_{u,\ell,i}(\vec{z}, \vec{w}, \vec{y}) \oplus \chi_{u,\ell,i}(\mathbf{z}^L, \vec{y})$  // Eq. (4)
11  $D \leftarrow \emptyset, R \leftarrow \{R_u \leftarrow \infty : u \in V\}$ 
12 for  $u \in V, i \in [\tau]$  do
13   if  $\zeta_{u,k,i}(\vec{z}, \vec{w}, \vec{y}) \neq 0$  then
14      $R_u \leftarrow \min(R_u, i)$  //  $R_{u,i} \leftarrow 1$  to obtain all reachable timestamps with
       space  $\mathcal{O}(n\tau)$ 
15     if  $u \notin D$  then
16        $D \leftarrow D \cup \{u\}$ 
17 return  $R, D$ 

```

Let us then turn to our algorithmic results.

Theorem 6 *There exists a randomized algorithm for solving k -RESTLESSMOTIFREACH problem in time $\mathcal{O}(2^k km \Delta)$ and.*

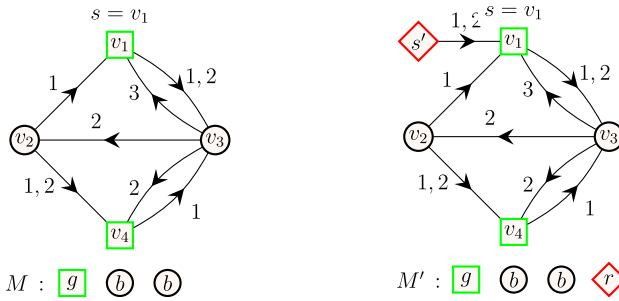


Fig. 4 An illustration of graph construction to solve k -RESTLESSMOTIFREACH. An instance of k -RESTLESSMOTIFREACH with a multiset of colors (left) and the transformation of the graph and the multiset (right). For illustrative purposes, we denote the vertices with color red, black and green using diamond, circular and square shapes, respectively

Proof Given an instance (G, c, k, M, s) of k -RESTLESSMOTIFREACH, we build a graph $G' = (V', E')$ such that $V = V \cup \{s'\}$, $E' = E \cup \{(s', s, i) : (s, u, i) \in E\}$, $c'(u) = c(u)$ for all $u \in V$, $c'(s') = k + 1$, $M' = M \cup \{c'(s')\}$ and query the FINEGRAINEDORACLE with instance $(G', c', k + 1, M')$. The construction is depicted in Fig. 4. In the instance $(G', c', k + 1, M')$, the origin of the graph is enforced by introducing an additional vertex s' adjacent to s . Since s' is assigned a unique color, if there is a resting path agreeing with M' , then the path must originate from s' and pass through s . If there exists a restless path originating from s' and ending at $u \in V \setminus \{s, s'\}$ such that the vertex colors of the path agree with M' , then we have a restless path originating from s and ending at u such that the vertex colors of the path agree with M . As the graph G' will have at most $2m$ edges and $n + 1$ vertices, we have obtained an algorithm for solving k -RESTLESSMOTIFREACH using $\mathcal{O}(2^k km \Delta)$ time and $\mathcal{O}(n \Delta)$ space. \square

An algorithm for k -RESTLESSREACH is obtained by transforming it to a k -RESTLESSMOTIFREACH instance. More precisely, the algorithm works by constructing a vertex-colored instance that encodes the source, whereas a multiset of colors encodes the path length and the solution to RESTLESSREACH following as a special case.

Theorem 7 *There exists a randomized algorithm for k -RESTLESSREACH in time $\mathcal{O}(2^k km \Delta)$ and.*

Proof Given an instance (G, s, k) of k -RESTLESSREACH, we introduce a coloring function $c : V \rightarrow \{1, 2\}$ such that $c(s) = 2$ and $c(v) = 1$ for all $v \in V \setminus \{s\}$. We obtain a graph $G' = (V, E \setminus \{(u, v, i) \in E : v = s\})$ by removing all incoming edges to s in G , and by setting the multiset $M = \{1^{k-1}\} \cup \{2\}$. We query the FINEGRAINEDORACLE with instance (G', c, k, M) . The transformation is illustrated in Fig. 5. In the instance (G', c, k, M) , the origin of the restless path is enforced by removing all incoming edges to s and coloring s with a unique color. If we have a restless path ending at $u \in V \setminus \{s\}$ and agreeing with multiset M , it implies that the temporal path originates from s and ends at u . The graph G' has $n + 1$ vertices and m edges, so we have a $\mathcal{O}(2^k km \Delta)$ time and $\mathcal{O}(n \Delta)$ space algorithm for solving k -RESTLESSREACH. \square

Theorem 8 *There exists a randomized algorithm for RESTLESSREACH in time $\mathcal{O}(2^n nm \Delta)$ and.*

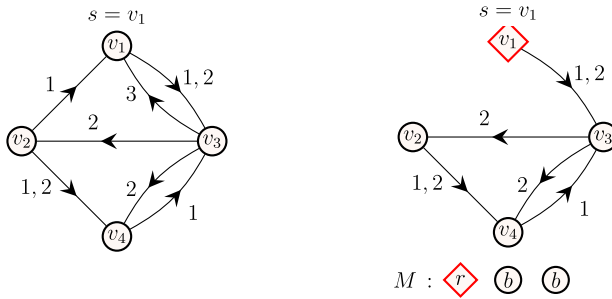


Fig. 5 The transformation of k -RESTLESSREACH instance (left) to k -RESTLESSMOTIFREACH instance (right). For illustrative purposes, we denote vertices with color red and black using diamond and circular shapes, respectively

Proof For solving RESTLESSREACH, the construction is similar to Theorem 7. However, we need to make $n - 2$ calls to the FINEGRAINEDORACLE assuming the maximum length of the restless path is $n - 1$. Finally, we obtain a set $R' = \{R'_{u,i} : u \in V, i \in [\tau]\}$ such that $R'_{u,i} = 1$ if there exists a restless path from s to u ending at time i with length at most $n - 1$, and $R'_{u,i} = 0$, otherwise. The pseudocode is available in Algorithm 2.

In total, we make $n - 2$ FINEGRAINEDORACLE calls for each $k \in \{2, \dots, n - 1\}$. Each run of the oracle takes $\mathcal{O}(2^k km \Delta)$ time. To summarize, the run time of the algorithm is $\sum_{k=2}^{n-1} 2^k km \Delta$, which is $\mathcal{O}(2^n nm \Delta)$. The space complexity is $\mathcal{O}(n\tau)$, completing the proof. \square

Algorithm 2: RESTLESSREACH($G = (V, E), s$)

Input: $G = (V, E)$, input graph
 s , source vertex
Output: R , minimum reachability time
 D , set of reachable vertices

```

1  $n \leftarrow |V|$ 
2 for  $u \in (V \setminus \{s\})$  do
3    $c(u) \leftarrow 1$ 
4  $c(s) \leftarrow 2$ 
5  $G' \leftarrow (V, E \setminus \{(u, s, i) \in E\})$ 
6 for  $k \in \{2, \dots, n - 1\}$  do
7    $M \leftarrow \{1^{k-1}\} \cup \{2\}$ 
8    $R, D \leftarrow \text{FINEGRAINEDORACLE}(G', c, k, M)$ 
9 return  $R, D$ 

```

4.6 Discussion

Consider the following variant of the restless reachability problem that we call at-most- k -restless reachability problem (*atm- k -RESTLESSREACH*). Given a temporal graph and a source vertex, we need to find a set of vertices which are reachable from source via a restless path such that the length of the path is at most $k - 1$. From Theorem 8, we have a randomized

algorithm for solving *atm*- k - RESTLESSREACH in time $\mathcal{O}(2^k km \Delta)$ and space $\mathcal{O}(n \Delta)$. Also note that we can use the algorithms for k - RESTLESSMOTIFREACH, k - RESTLESSREACH and RESTLESSREACH as we can solve k - RESTLESSMOTIF, k - RESTLESSPATH and RESTLESSPATH, respectively.

A general variant of RESTLESSREACH with a set of sources $S \subseteq V$ can be reduced to RESTLESSREACH with a single source by introducing an additional vertex s' and connecting all the sources $s \in S$ to s' with a temporal edge. More precisely, given a graph $G = (V, E)$ and set of sources $S \subseteq V$, we construct a graph $G' = (V', E')$ with $V' = V \cup \{s'\}$ and $E' = E \cup \{(s', s, i) \mid s \in S \text{ and } (s, u, i) \in E\}$. Solving RESTLESSREACH on the graph instance G' with source s' is equivalent to solving RESTLESSREACH with set of sources S . Additionally, finding the restless path that minimizes the length (shortest path), minimizes the arrival time (fastest path), or minimizes the total waiting time (foremost path) can be computed using the output $\{R'_{u,i} : u \in V, i \in [\tau]\}$ for each $\ell \in \{2, \dots, k\}$ from the fine-grained oracle. However, enabling such computation requires $\mathcal{O}(n \tau k)$ space.

4.7 Extracting an optimal solution using k queries

In this section, we present an algorithm for extracting an optimal solution for k - RESTLESSMOTIF and k - RESTLESSPATH using k queries to the FINEGRAINEDORACLE. By optimal we mean that the maximum timestamp in the restless path is minimized.

Our FINEGRAINEDORACLE reports the existence of a restless path from a given source to a destination at discrete timestamps with a YES/NO answer. However, in many cases we also require an explicit solution, i.e., a restless path which actually witnesses the fact. We present two approaches to extracting an optimal solution. Our first approach makes use of self-reducibility of decision oracles and temporal DFS, based on the previous work of Björklund et al. [8] and Thejaswi et al. [54]. Our second approach makes use of the fine-grained oracle. For using self-reducibility, we implement a naive version of the multilinear sieve which only reports a YES/NO answer without the fine-grained capabilities, that is, the algorithm does not report the set of vertices and the timestamps at which the vertices are reachable from a given source via a restless path. We summarize our results in Table 1.

Self-reducibility and temporal DFS. The approach works in three steps: First, we obtain the minimum (optimal) timestamp $t \in [\tau]$ for which there exists a feasible solution. For this, we construct the polynomial encoding of restless walks of length $k - 1$ which end at time at most $t' \in [\tau]$ and query the decision oracle for the existence of a solution. Using binary search on the range $[\tau]$, we use at most $\log \tau$ queries to obtain the optimal timestamp. Next, we extract a k -vertex temporal subgraph that contains a restless temporal path. By recursively dividing the graph into two halves, we can obtain the desired subgraph using $\mathcal{O}(k \log n)$ queries to the decision oracle in expectation [8]. Finally, we extract the restless path by performing a temporal DFS from a given source in the k -vertex subgraph. Even though the worst-case complexity of the temporal DFS is $\mathcal{O}(k!)$, we demonstrate that the approach is practical. For technical details of self-reducibility of decision oracles, we refer the interested reader to Björklund et al. [8].

In summary, the overall complexity of extracting an optimal solution using self-reducibility and temporal DFS is $\mathcal{O}((2^k k^2 m \Delta \log n \log \tau) + k!)$.

Using fine-grained oracle. As a second approach, extracting a solution can also be done with k queries to the fine-grained oracle. Let us present the high-level idea behind fine-grained extraction. Consider a temporal graph presented in Fig. 6 with vertex set $V = \{v_1, \dots, v_6\}$

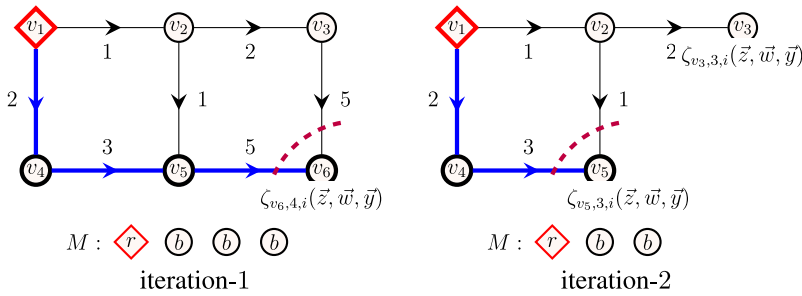


Fig. 6 Extracting a restless path using the fine-grained decision oracle. On iteration one (left), a k -RESTLESS-PATH instance with a multiset of colors M and a restless path of length 3 from vertex v_1 to vertex v_6 highlighted in bold. On iteration two (right), a k -RESTLESSPATH instance with a multiset of colors M and a restless path of length 2 from vertex v_1 to vertex v_5 highlighted in bold

with resting times $\delta(v_1) = \dots = \delta(v_6) = 2$. In this example, we want to extract a restless path of length 3 from vertex v_1 to vertex v_6 , if such a restless path exists. For illustrative purposes, we use colors **black** and **red** and have the multiset of colors $M = \{r, b, b, b\}$. For the first iteration, it suffices to verify if $\zeta_{v_6,4,i}(\vec{z}, \vec{w}, \vec{y})$ evaluates to a nonzero term for each $i \in [\tau]$. Since there exists a restless path of length 3 ending at vertex v_6 at time 5 agreeing with the colors in M , the corresponding evaluation polynomial $\zeta_{v_6,4,5}(\vec{z}, \vec{w}, \vec{y})$ is nonzero. For the second iteration, delete vertex v_6 from the graph and remove a b from M leaving us with $M = \{r, b, b\}$. Now check if there exists a restless path of length 2 ending at any of the neighbors of v_6 , i.e., $N_5(v_6) = \{v_3, v_5\}$ at any of the timestamps $i \in \{3, 4, 5\}$. This can be done by verifying if the polynomials $\zeta_{v_3,2,i}(\vec{z}, \vec{w}, \vec{y}), \zeta_{v_5,2,i}(\vec{z}, \vec{w}, \vec{y})$ evaluate to a nonzero term for timestamps $i \in \{3, 4, 5\}$. In our case, $\zeta_{v_5,2,3}(\vec{z}, \vec{w}, \vec{y})$ evaluates to a nonzero term, which implies that there exists a restless path from v_1 to v_3 ending at timestamp 3 agreeing with the colors in M' , so we add the edge $(v_5, v_6, 5)$ to the solution. We can recursively repeat the second iteration until we reach the vertex v_1 to obtain a restless path from v_1 to v_5 .

A generalization of the approach is described as follows: Let (G, c, M, k, s, d, τ) be an instance of k -RESTLESSMOTIF. We build an instance $(G^\ell, c', M^\ell, \ell, s', \tau^\ell)$ of k -RESTLESSMOTIFREACH for each $\ell \in \{k + 1, k, \dots, 2\}$. For the first iteration, where $\ell = k + 1$, the graph G^ℓ is constructed as described in Theorem 6 to obtain a new source vertex s' and a coloring function $c', M^\ell = M \cup c'(s')$, and $\tau^\ell = \tau$. The graph construction of k -RESTLESSMOTIFREACH for $\ell = k + 1$ is illustrated in Fig. 7. We apply the algorithm from Theorem 6 to obtain R^ℓ and D^ℓ .

Let $(u_\ell, i_\ell) \in R^\ell$ be a (vertex, minimum timestamp) pair such that $R_{u_\ell, i_\ell}^\ell = 1$. For the first iteration, where $\ell = k + 1$, we remove the vertex $u_{k+1} = d$ and remove the incoming and outgoing edges of u_{k+1} in G^{k+1} to obtain G^k . Let $M^{\ell-1} = M^\ell \setminus (u_\ell)$ and $\tau^{\ell-1} = i_\ell$. We evaluate the instance $(G^{\ell-1}, c', M^{\ell-1}, \ell - 1, s', \tau^{\ell-1})$ to obtain $R^{\ell-1}$. Let $(u_{\ell-1}, i_{\ell-1}) \in R^{\ell-1}$ be a (vertex, timestamp) pair such that $R_{u_{\ell-1}, i_{\ell-1}}^{\ell-1} = 1$. As $R_{u_\ell, i_\ell}^\ell = 1$ there exists an edge $(u_{\ell-1}, u_\ell, i_\ell) \in E(G^\ell)$ in G^ℓ . In each iteration, we add the edge $(u_{\ell-1}, u_\ell, i_\ell)$ to the solution and continue the process for each $\ell \in \{k + 1, k, \dots, 2\}$. In total we make k queries to the fine-grained oracle. Thus, the run time of extracting an optimal solution is $\sum_{\ell=2}^{k+1} 2^\ell \ell m \Delta = \mathcal{O}(2^k k m \Delta)$. Similarly, we have a $\mathcal{O}(2^k k m \Delta)$ -time and algorithm for extracting an optimal solution for k -RESTLESSPATH.

Using a similar construction, we can improve the run time of extracting an optimal solution for PATHMOTIF and k -TEMPPATH introduced in [53, 54] from $\mathcal{O}((2^k k (n\tau + m))(k \log n +$

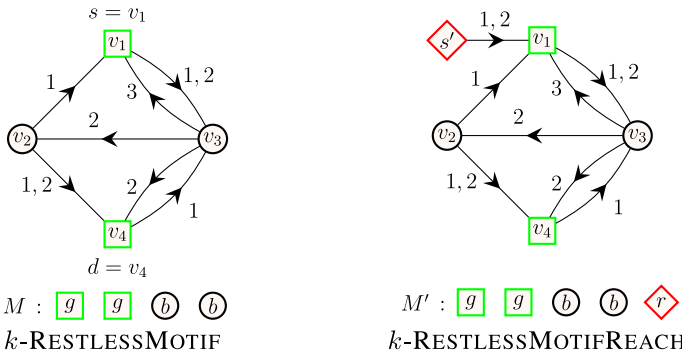


Fig. 7 The transformation of k -RESTLESSMOTIF (left) to an instance of k -RESTLESSMOTIFREACH (right) for extracting a solution using the fine-grained oracle. Vertices with color red, black and green are drawn using diamond, circular and square shapes, respectively

$\log \tau) + k!)$ to $\mathcal{O}(2^k k(n\tau + m))$. Additionally, our fine-grained construction can be employed to extract a solution for the k -path problem and the graph motif problem in static graphs by reducing the number of queries from $\mathcal{O}(k \log n)$ to $\mathcal{O}(k)$, thus improving the work of Björklund et al. [8].

4.8 A deterministic $\mathcal{O}^*(2^n)$ algorithm for solving RESTLESSPATH

To obtain a deterministic $\mathcal{O}^*(2^n)$ -time algorithm for RESTLESSPATH, we proceed as follows. First, we construct a static expansion of the given temporal graph and then execute a deterministic algorithm for RAINBOWPATH on the resulting graph. Before proceeding, we introduce this problem formally.

Rainbow path problem in static graphs (RAINBOWPATH). Given a static graph $G' = (V', E')$, a vertex-coloring function $c : V' \rightarrow C'$ and two distinct vertices $s, d \in V'$, the problem asks us to decide if there exists a rainbow path from s to d , that is, a path on which no color repeats. A deterministic polynomial-space algorithm for RAINBOWPATH is due to Kowalik and Lauri [36].

Lemma 9 ([36], Corollary 5) *There exists a deterministic algorithm for solving RAINBOWPATH in static graphs with time $\mathcal{O}(2^{|C'|} |C'|(|E'| + |C'|^2))$ and space $\mathcal{O}(|V'| + |C'|)$, where V' is the set of vertices, E' is the set of edges and C' is the set of vertex colors.*

δ -expansion. Let $G = (V, E)$ be a temporal graph with maximum timestamp τ with number of vertices $|V| = n$ and number of edges $|E| = m$. For simplicity, we assume that the vertex set is $V = \{v_1, \dots, v_n\}$. Let $\delta : V \rightarrow \mathbb{N}_+$ be a vertex-dependent waiting time. The δ -expansion of G is a static directed graph $G^\uparrow(\delta) = (V', E')$ where $V' = \{v_i^t : v_i \in V, t \in [\tau] \mid \exists e \in E : v_i \in e\}$ and $E' = \{(v_i^t, v_j^{t+\ell}) : (v_i, v_j, t) \in E, \ell \in \{0, \dots, \delta(v_j)\}, t + \ell \leq \tau\}$.⁸ For each vertex $v_i^t \in V'$, we assign the color $c(v_i^t) = i$ and the set of colors $C = [n]$. Observe that $|V'| = |E'| = (\Delta + 1)|E|$, where $\Delta = \max_{v_i \in V} \delta(v_i)$. Note that $G^\uparrow(\delta)$ can be computed in time and space $\mathcal{O}(m\Delta)$. For an illustration of δ -expansion, see Fig. 8. Finally, we claim that there exists a RESTLESSPATH in G if and only if there exists a RAINBOWPATH in $G^\uparrow(\delta)$.

⁸ Note that V' is not a multiset, that is, V' does not contain duplicate elements.

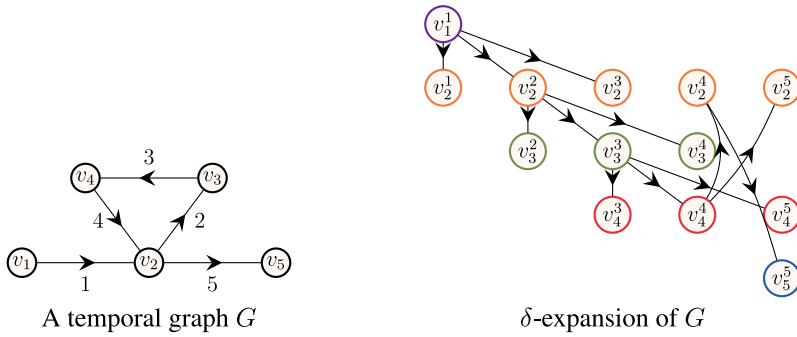


Fig. 8 On the left, a temporal graph $G = (V, E, \tau)$ with vertex set $V = \{v_1, \dots, v_5\}$, edge set $E = \{(v_1, v_2, 1), (v_2, v_3, 2), (v_3, v_4, 3), (v_4, v_2, 4), (v_2, v_5, 5)\}$, maximum timestamp $\tau = 5$ and resting time $\delta(v_1) = \dots = \delta(v_5) = 2$. On the right, δ -expansion $G^\uparrow(\delta)$ of G . For each temporal edge $(v_i, v_j, t) \in E$, we introduce at most $\delta(v_j) + 1$ static edges $\{(v_i^t, v_j^{t+\ell}) : \ell \in \{0, \dots, \delta(v_j)\}, t + \ell \leq \tau\}$. If $(v_i^t, v_j^t) \in E'$, then $v_i^t, v_j^t \in V'$. Finally, all vertices with the same subscript, i.e., $\{v_i^1, v_i^2, \dots\}$ for each $i \in [n]$, are assigned the same color

Lemma 10 *There exists a RESTLESSPATH in the temporal graph G if and only if there exists a RAINBOWPATH in the corresponding δ -expansion $G^\uparrow(\delta)$.*

Proof Let $P = v_1 e_1 v_2, \dots, e_{k-1} v_k$ be a restless path in G . We construct a rainbow path $P' = v'_1 e'_1 v'_2 \dots e'_{k-1} v'_k$ as follows. For edges $e_i = (v_i, v_{i+1}, t_i), e_{i+1} = (v_{i+1}, v_{i+2}, t_{i+1}) \in P, i = [k - 1]$, we pick a static edge $e'_i = (v_i^{t_i}, v_{i+1}^{t_{i+1}})$ to path P' and vertices $v'_i = v_i^{t_i}, v'_{i+1} = v_{i+1}^{t_{i+1}}$. From construction, we know that $v'_i = v_i t_i, i \in [k]$ and $e'_i = (v_i^{t_i}, v_{i+1}^{t_{i+1}}), i \in [k - 1]$ exists in $G^\uparrow(\delta)$. Since P is a path, no vertex repeats in P , which implies that no two vertices in P' have the same color. We conclude that P' is a rainbow path.

Conversely, let $P' = v'_1 e'_1 v'_2 \dots e'_{k-1} v'_k$ be a rainbow path in $G^\uparrow(\delta)$, we construct a restless path $P = v_1 e_1 v_2 \dots e_{k-1} v_k$ by replacing each edge $e'_i = (v_i^{t_i}, v_{i+1}^{t_{i+1}}) \in P', i = [k - 1]$ by $e_i = (v_i, v_{i+1}, t_i)$. From construction the waiting time at vertex v_i is at most $t_{i+1} - t_i \leq \delta(v_i)$. Since P' is a rainbow path, it implies that $v_i \neq v_j$ for all $i \neq j$. So P is a restless path, which concludes our proof. \square

Combining Lemma 9 with Lemma 10, we obtain the following result.

Theorem 11 *There exists a deterministic algorithm for solving the RESTLESSPATH problem in time $\mathcal{O}(2^n n(m\Delta + n^2))$ and space $\mathcal{O}(m\Delta + n)$.*

Theorem 11 also implies a deterministic $\mathcal{O}(2^n n^2(m\Delta + n^2))$ -time algorithm for solving RESTLESSREACH. However, the deterministic algorithm presented in this section is of theoretical interest only and is unlikely to scale in practice even for graphs of modest size. However, note that Lemma 9 can be replaced with any algorithm for RAINBOWPATH. Indeed, any improvement in the running time of the algorithm for RAINBOWPATH will improve the running time of algorithms for RESTLESSPATH and RESTLESSREACH.

4.9 Infeasibility of a $\mathcal{O}^*((2 - \epsilon)^k)$ -time algorithm for k - RESTLESSMOTIF

In this section, we prove that under plausible complexity-theoretic assumptions, the algorithms presented for k - RESTLESSMOTIF and k - RESTLESSMOTIFREACH are optimal.

Table 1 A summary of time and space complexity

Problem	Time complexity	Space complexity
<i>Fine-grained oracle (randomized)</i>		
RESTLESSPATH	$\mathcal{O}(2^n nm \Delta)$	$\mathcal{O}(n \Delta)$
k - RESTLESSPATH	$\mathcal{O}(2^k km \Delta)$	$\mathcal{O}(n \Delta)$
k- RESTLESSMOTIF	$\mathcal{O}(2^k km \Delta)$	$\mathcal{O}(n \Delta)$
RESTLESSREACH	$\mathcal{O}(2^n nm \Delta)$	$\mathcal{O}(n \Delta)$
k - RESTLESSREACH	$\mathcal{O}(2^k km \Delta)$	$\mathcal{O}(n \Delta)$
k- RESTLESSMOTIFREACH	$\mathcal{O}(2^k km \Delta)$	$\mathcal{O}(n \Delta)$
atm- k - RESTLESSREACH	$\mathcal{O}(2^k km \Delta)$	$\mathcal{O}(n \Delta)$
<i>Inclusion–exclusion (deterministic)</i>		
RESTLESSPATH	$\mathcal{O}(2^n n(m \Delta + n^2))$	$\mathcal{O}(m \Delta + n)$
RESTLESSREACH	$\mathcal{O}(2^n n^2(m \Delta + n^2))$	$\mathcal{O}(m \Delta + n)$
<i>Extraction (k - RESTLESSPATH)</i>		
Self-reducibility + temporal DFS	$\mathcal{O}((2^k k^2 m \Delta \log n \log \tau) + k!)$	$\mathcal{O}(n \Delta)$
Fine-grained extraction	$\mathcal{O}(2^k km \Delta)$	$\mathcal{O}(n \tau)$

Here, n is the number of vertices, m is the number of edges, τ is the maximum timestamp, $k - 1$ is the length of path and Δ is the maximum resting time. For extraction we use randomized algorithm as a subroutine

In particular, the assumption that we rely on is the Set Cover Conjecture [18] (SCC), which is formulated as follows. In the SETCOVER problem, we are given an integer f and a family of sets \mathcal{S} over the universe $U = \bigcup \mathcal{S}$ with $n = |U|$ and $m = |\mathcal{S}|$. The goal is to decide whether there is a subfamily of at most f sets $S_1, S_2, \dots, S_f \in \mathcal{S}$ such that $U = \bigcup_{i=1}^f S_i$, i.e., that the selected sets cover the universe U . The SCC of Cygan et al. [18] states that there is no algorithm for the SETCOVER problem that runs in time $(2 - \epsilon)^n (nm)^{\mathcal{O}(1)}$ for any $\epsilon > 0$. In fact, the fastest known algorithm for solving SETCOVER runs in time $\mathcal{O}^*(2^n)$, and an algorithm running in time $\mathcal{O}^*((2 - \epsilon)^n)$ for any $\epsilon > 0$ has been deemed a major breakthrough after decades of research on the problem. Under SCC, exponential lower bounds for several fundamental problems are known (see e.g., [6, 18, 37]). However, even if SCC turned out to be false, these results (and Theorem 13) are still meaningful: instead of trying to find a faster algorithm for e.g., the arguably richer problem of k- RESTLESSMOTIF, one can focus on SETCOVER which is simpler.

To obtain our result, it is convenient to recall the COLORFULPATH problem in static graphs and to perform a reduction from that problem to k- RESTLESSMOTIF problem.

Colorful path problem (COLORFULPATH). Given a static graph $G = (V, E)$ and a coloring function $c : V \rightarrow [k]$, the problem asks if there exists a path of length $k - 1$ in G such that the vertex colors of the path are different (i.e., such that each color occurs exactly once). The problem is known to be NP-hard, and known not to admit a $\mathcal{O}^*((2 - \epsilon)^k)$ algorithm for any $\epsilon > 0$ assuming SCC.

Theorem 12 (Kowalik and Lauri [36]) *Assuming the Set Cover Conjecture, there exists no $\mathcal{O}^*((2 - \epsilon)^k)$ time algorithm for solving the COLORFULPATH problem for any $\epsilon > 0$.*

Theorem 13 *If the k- RESTLESSMOTIF problem has a $\mathcal{O}^*((2 - \epsilon)^k)$ time algorithm for any $\epsilon > 0$ then COLORFULPATH problem has a $\mathcal{O}^*((2 - \epsilon)^k)$ time algorithm.*

Proof Given an instance $I = (G, c, k)$ of COLORFULPATH in static graphs, we construct an instance $I' = (G', M', c', k')$ of k- RESTLESSMOTIF in temporal graph by letting $G' = (V', E')$, $V' = V \cup \{s, d\}$, $E' = \{(u, v, 1) : (u, v) \in E\} \cup \{(s, u, 1) : u \in V\} \cup \{(u, d, 1) : u \in V\}$, $c'(u) = c(u)$ for all $u \in V$, $c(s) = k + 1$, $c(d) = k + 2$, $M = \{1, \dots, k + 2\}$, and $\delta : V \rightarrow 1$, $\Delta = 1$, $k' = k + 2$. Informally, G' is constructed from G by replacing each edge with a temporal edge with timestamp one, and by making s and d adjacent to a new vertex both receiving a new unique color. We claim that the instance I of COLORFULPATH has a solution if and only if the instance I' of k- RESTLESSMOTIF has a solution.

Let $P = v_1e_1v_2 \dots e_{k-1}v_k$ be a colorful path in G . By construction, we know that for each edge $(u, v) \in E$ we have $(u, v, 1) \in E'$, so the path $P' = se'_0v_1e'_1v_2 \dots e'_{k-1}v_ke'_kd$ exists in G' , where $e'_0 = (s, v_1, 1)$, $e'_k = (v_k, d, 1)$, and $e'_i = (v_i, v_{i+1}, 1)$ for all $i \in [k - 1]$. Also, the vertex colors of $\{v_1, \dots, v_k\}$ agree with $\{1, \dots, k\}$ since the P is colorful, so the vertex colors of $\{s, d\} \cup \{v_1, \dots, v_k\}$ agree with M' . We conclude that P' is a k- RESTLESSMOTIF in G' . Conversely, let $P' = se'_0v_1e'_1 \dots e'_{k-1}v_ke'_{k+1}d$ be a solution for I' in G' . We construct a static path $P = v_1e_1v_2 \dots e_{k-1}v_k$ by replacing the edges $e'_i = (v_i, v_{i+1}, 1)$ by $e_i = (v_i, v_{i+1})$ for all $i \in [k + 1]$. Since the vertices $\{s, d\} \cup \{v_1, \dots, v_k\}$ agree with colors $M = \{1, \dots, k + 2\}$, the vertex colors of $\{v_1, \dots, v_k\}$ agree with colors $\{1, \dots, k\}$ as the colors $k + 1$ and $k + 2$ only appear once each on G' . Evidently, P is a colorful path in G . □

It follows that if we have an algorithm for solving k- RESTLESSMOTIF with $\mathcal{O}^*((2 - \epsilon)^k)$ time for some $\epsilon > 0$, we can use it solve COLORFULPATH in static graphs using the construction described in Theorem 13 within the same time bound. However, from Theorem 12 we know that such an algorithm does not exist unless SCC is false. Put differently, assuming SCC, k- RESTLESSMOTIF does not admit an algorithm running in time $\mathcal{O}^*((2 - \epsilon)^k)$ for any $\epsilon > 0$. Finally, since k- RESTLESSMOTIFREACH problem generalizes the k- RESTLESSMOTIF problem, the former problem does not admit an algorithm running in time $\mathcal{O}^*((2 - \epsilon)^k)$ for any $\epsilon > 0$, assuming SCC.

5 Experiments

In this section, we describe our setup and the experimental results to validate our approach and demonstrate its scalability. Our implementation is available as open source [55].

5.1 Implementation

A high-level intuition of the implementation is as follows: For variables $\vec{x} = \{x_v : v \in V\}$ and $\vec{y} = \{y_{uv,\ell,i} : (u, v, i) \in E, \ell \in [k]\}$ we assign a value from the Galois field $\text{GF}(2^b)$. Multiplication between any two field variables is defined as an XOR operation, likewise, multiplying two variables with same value results in a zero-term. We know that the monomials corresponding to walks that are not paths have at least one repeated variable, so the corresponding monomial evaluates to a zero term, while the monomial corresponding to a path evaluates to a nonzero term since there are no repeated variables. It is possible that a monomial corresponding to a path might evaluate to a zero term resulting in a false negative. To reduce the probability of false negatives, we repeat the evaluation with 2^k random assignments for variables $\{x_u : u \in V\}$ and $\{y_{uv,\ell,i} : (u, v, i) \in E, \ell \in [k]\}$. In theory, the false-negative probability of our algorithm is $2^{-b}(2k - 1)$. For our experiments, we choose the field size $b = 64$, which makes the false-negative probability negligible.

Modern CPUs have very high arithmetic and memory bandwidth; however, the bandwidth comes at the cost of latency. Each arithmetic and memory-access operation is associated with a corresponding latency factor, and often the memory-access latency is orders of magnitude greater than the arithmetic latency. As such, the challenge for efficient implementation engineering is to keep the arithmetic pipeline busy while fetching data from memory for the subsequent arithmetic operations. Memory bandwidth can be improved by using coalesced memory access, that is, by organizing the memory layout such that the data for consecutive computations are available in consecutive memory addresses. In addition, we can use hardware prefetching to fetch the data required in subsequent computation while performing computation on the data, which is currently in the memory. Arithmetic bandwidth can be improved using vector extensions, that is, by grouping the data on, which the same arithmetic operations are executed. More precisely, if we are executing the same arithmetic instructions on different operands or data, we can group the operands using vector extensions to execute arithmetic operations in parallel. For more technical details related to implementation engineering, we refer the reader to [10, 32, 54].

Our implementation is written in the C programming language with OpenMP constructs to achieve thread-level parallelism. Vector parallelism is achieved by enabling parallel executions of the same arithmetic operations, which make use of advanced vector extensions (AVX2). Additionally, we use carry-less multiplication of one quadword (PCLMULQDQ) instruction set to enable fast finite-field arithmetic. The finite-field arithmetic implementation we use is from [10].

Our engineering effort boils down to implementing the recursions in Eqs. (1) and (3) and evaluating the polynomial using 2^ℓ random substitutions for the x -variables. Recall that from the construction of the generating function the y -variables are unique, so we generate the values of y -variables using a pseudorandom number generator. The values of the x -variables are computed using Eq. (3). Our implementation loops over four variables: the outer most loop is over $[\ell]$, the second loop is over $[\tau]$, the third loop is over V , and the final loop is over $\{0, \dots, \delta(u)\}$ for $u \in V$. In Eq. (3), computing the polynomial $\chi_{u,\ell,i}(\vec{x}, \vec{y})$ is independent for each $u \in V$ if we fix ℓ and i , so the algorithm can be thread-parallelized up to $|V| = n$ threads. We make use of the OpenMP API using the `omp parallel for` construct with `default` scheduling over vertices in V to achieve thread parallelism. Additionally, performing 2^ℓ random substitutions of x -variables is independent of each other, so each of the 2^ℓ evaluations can be vector-parallelized. We achieve this by grouping the arithmetic operations on 2^ℓ random substitutions of variables in x and enabling the vector extensions from AVX2. Recall that our inner-most loop is over $\{0, \dots, \delta(u)\}$, so we arrange the memory layout as $n \times \tau$ to saturate the memory bandwidth. We also employ hardware prefetching by forcing the processor to fetch data for subsequent computations while we are still performing the computation on the data, which is already in the memory.

Our implementation uses $\mathcal{O}(n\tau + m)$ memory, which is due to the adjacency list representation of the temporal graphs.

Preprocessing. In the restless reachability problems considered in our work, we compute reachability from a given source vertex to all other vertices without an explicit restriction on the time window. As such, we do not see a straightforward approach to preprocessing the temporal graph G to reduce its size using heuristic preprocessing techniques such as slicing G within a time window, i.e., considering the edges between a minimum and maximum timestamp window. Alternatively, we can merge G to obtain a static graph, compute reachability on the static graph, and reconstruct a temporal graph by only using the vertices, which are reachable in the static graph. Such a preprocessing technique is correct since there

exists a restless path between any two vertices in a temporal graph only if there exists a path in the corresponding static graph, while the other direction is not always true. However, most of the datasets considered in our experiments have a connected static underlying graph, and therefore we do not see a significant reduction in graph size.

For the restless reachability problems with additional color constraints i.e., for the k -RESTLESSMOTIF and k -RESTLESSMOTIFREACH problems, we can take advantage of two preprocessing techniques to reduce the graph size: (i) by removing all the vertices whose vertex colors do not match with the multiset colors; (ii) by merging the temporal graph to a static graph instance, build a vertex-localized sieve on the static graph, and reconstruct the graph using the set of vertices reachable in the static graph instance. Note that these are heuristic approaches to reduce the graph size and we do not claim any theoretical bounds for the reduction in the graph size. For a detailed discussion of preprocessing using vertex-localized sieving we refer the reader to an earlier work [54, § Preprocessing].

5.2 Experimental setup

Here we describe the hardware details and the input graphs used for our experiments.

Hardware. We make use of two hardware configurations for our experiments.

- A *workstation* with 1×3.2 GHz Intel Core i5-4570 CPU, Haswell microarchitecture, 4 cores, 32 Gb memory, Ubuntu, and gcc v9.1.0.
- A *computenode* with 2×2.5 GHz Intel Xeon 2680 V3 CPU, 24 cores, 12 cores/CPU, 256 Gb memory, Red Hat, and gcc v9.2.0.

The experiments make use of all the cores. Additionally, we make use of AVX2 and hardware prefetching to saturate the arithmetic and memory bandwidth, respectively.

Input graphs. We use both synthetic and real-world graphs in our experiments. For *synthetic* graphs, we use the temporal graph generator from Thejaswi et al. [54, § 9.3], in particular we make use of d -regular and power-law graphs. The regular graphs are generated using the *configuration model* [11, § 2.4]. The configuration model for power-law graphs is as follows: given non-negative integers D , n , w , and $\alpha < 0$, we generate an n -vertex graph such the following properties roughly hold: (i) the sum of vertex degrees is Dn ; (ii) the distribution of degrees is supported at w distinct values with geometric spacing; and (iii) the frequency of vertices with degree d is proportional to d^α . The edge timestamps are assigned uniformly at random in the range $[\tau]$. Both directed and undirected graphs are generated using the same configuration model; however, for directed graphs the orientation is preserved. We ensure that the graph generator produces identical graph instances in all the hardware configurations.

For *real-world* graphs, we use the co-presence dataset from socio-patterns [23], Koblenz network collections [40], Copenhagen study network [49], and public-transport networks [39]. For a description of the datasets, see the respective references. The preprocessing details of each dataset are described below. We preprocess the datasets to generate a graph by renaming the location identifiers (or vertices) in the range from 1 to the maximum number of locations (or vertices) available in the dataset. If the time values in dataset are Unix timestamps, we approximate the value to the closest second rounded down before assigning a unique timestamp identifier. For datasets from socio-patterns, we reduce the maximum timestamp τ by dividing each timestamp by 20 and rename the timestamps in the range from 1 to the difference between the maximum and minimum timestamps. Since socio-patterns datasets are undirected contact-networks we replace each undirected edge with two directed edges in

both directions. For the Koblenz datasets we round the timestamp to the closest day. Finally, we present statistics of all datasets used for experiments in Appendix D.

5.3 Baselines

In this section, we discuss the baseline approach considered for comparison. Additionally, we present careful justification on why certain approaches used to solve temporal reachability problems fail to solve restless reachability problems in Appendix E.

Exhaustive search (baseline). We consider an exhaustive search algorithm based on temporal depth-first-search (DFS), which is a parameterized algorithm with respect to the maximum degree of the graph d_{\max} . We perform temporal DFS starting from a source s by respecting waiting-time constraints and restrict the depth of the search to k . We report the minimum reachability time for the vertices that are reachable from s by at most $k - 1$ hops. The time complexity of the exhaustive search algorithm is $\mathcal{O}(d_{\max}^k)$. As demonstrated in previous work [53, 54], and also shown in our experiments (see Sect. 5.4), the exhaustive search does not scale for large scale-free graphs. In particular, many real-world graphs exhibit a power-law degree distribution, or more generally, scale-free structure. However, the exhaustive search algorithm is highly practical for some structured graphs such as graphs that are close to being d -regular with small maximum degree. In our experiments, we refer to the exhaustive search algorithm as the *baseline*. Furthermore, we implement the FPT(k)-algorithm by Casteigts et al. [13], which uses a different polynomial construction, with running time $\mathcal{O}(2^k(kn + km\Delta))$ and space complexity $\mathcal{O}(kn\tau)$.⁹

Discussion. We stress that our focus is on exact computation of restless reachability. As such, we do not compare our proposed algorithms to probabilistic methods, heuristics or approximation algorithms. Further, to the best of our knowledge, there is no publicly available implementation for solving the restless reachability problem that scales to large graphs such as those considered in our experiments. For instance, we argue that the algorithms presented by [13] are unlikely to perform adequately in practice given their exponential dependency on a structural parameter that, in many real-world networks, has a high value (see Appendix D Table 7). In addition, it is not immediate how index computation techniques such as 2-hop or 3-hop covers or location-based indexing would extend to solve restless reachability.

Finally, a careful reader might question our focus on *exact computation* by recalling that our algorithm has a false-negative probability of $(2k - 1)/2^b$. However, by fixing a suitable value for b and potentially running our method multiple times, we can make this probability arbitrarily close to zero. For concreteness, we choose $b = 64$ for our experiments, which means that when say $k = 10$, the per-vertex false-negative probability is less than $2^{59} \approx 5.76 \cdot 10^{-17}$. In comparison, a modern consumer CPU running for at least five days has a 1 in 330 chance of a hardware failure due to a machine-check exception, a 1 in 470 chance of a disk subsystem failure, and 1 in 2700 ($\approx 3.7 \cdot 10^{-4}$) chance of a DRAM memory failure [45, Figure 2], all significantly more likely than our algorithm making an error.

5.4 Experimental results

In this section, we report our experimental results. The experiments are designed to study the following aspects:

⁹ The exact polynomial factors and space complexity are not detailed in their paper; the reported complexities are based on our implementation.

- (i) Scalability of the algorithm to graphs with up to 10 million edges on the workstation configuration;
- (ii) Scalability of the algorithm to large graphs with up to one billion edges on the computenode configuration;
- (iii) Computing restless reachability in real-world datasets; and
- (iv) A case study investigating the effectiveness of different vertex-selection strategies to act as barriers and minimize the spread of diffusion processes (e.g., infectious diseases) through the temporal network. An overview of our experiments is available in Table 2.

In our experimental results, run time refers to the empirical running time and memory refers to the peak-memory usage of our implementation.

Scalability. To demonstrate the scalability of the algorithm, we experiment with synthetic graphs. The experiments are performed on five independent random power-law graphs for each configuration specified in Table 2. The source vertex is chosen uniformly at random. All experiments are executed on the workstation configuration using all cores with directed graphs. Note that we demonstrated the scalability of the algorithm for k - RESTLESSREACH instances, however for other problem instances the run times are similar.

In Fig. 9 (left), we compare the run time as a function of the number of edges for the fine-grained oracle and the baseline. We observe a linear increase in the run time with the number of edges, as predicted by the theoretical analysis. In Fig. 9 (center-left), we compare the run time as a function of the length of the restless path for the fine-grained oracle and the baseline. We observe an exponential increase in the run time with the increase in length of the restless path, as predicted by the theoretical analysis. The variance in run time between the independent graph inputs is very small for the algebraic algorithm as compared to the baseline, which exhibit high variance in run time. We observe that our baseline fails to report the solution for power-law graphs with $n = 10^3$, $m = 10^6$, and $k = 10$. Note that we terminate experiments that take more than ten hours. The run time of the exhaustive search algorithm depends on the degree distribution of the graph: if the temporal DFS visits a high-degree vertex then the baseline algorithm takes long time to complete the execution. However, the baseline is efficient for sparse d -regular graphs, where the maximum degree is small. Again, for scaling with respect to the length of the restless path our baseline failed to report a solution for $k \geq 10$. We terminate the experiments, which take more than ten hours of run time. Our implementation can handle the k - RESTLESSREACH problem on a graph instance with one million nodes, ten million edges, and path length $k = 10$, in less than 20 min using less than 2 gigabytes of working memory on the workstation configuration.

In Fig. 9 (center-right), we report the run time of the algorithm as a function of the maximum resting time Δ . We do not observe a linear-scaling of the run time as the theoretical analysis tells us. A possible explanation is that the graphs used for the experiments are sparse, and there is not enough workload to saturate the empirical arithmetic and memory bandwidth of the hardware, simultaneously. However, with the increase in the resting time Δ we need to perform more arithmetic operations there by improving the arithmetic and memory bandwidth. This is due to the fact that the implementation enables more streamlining of the memory and arithmetic pipeline of the computer hardware when there is enough workload to parallelize.

Our final scalability experiments report the effect of the graph density on scalability. In Fig. 9 (right), we report the run time of the algorithm as a function of the graph density. Density of the graph is the ratio of number of edges and the number of vertices. We observe that our implementation performs better for dense graphs. Again a possible explanation is that

Table 2 Overview of the experiments

Experiment	n	m	d	k	Δ	δ	Sep (%)
<i>Figure 9</i>							
Left	$10^2, \dots, 10^5$	$10^4, \dots, 10^7$	10^2	10	10	$V \rightarrow \{10\}$	-
Center-left	10^3	10^5	10^2	$5, \dots, 15$	10	$V \rightarrow \{10\}$	-
Center-right	10^5	10^7	10^2	10	$10, 20, \dots, 100$	$V \rightarrow \{\Delta\}$	-
Right	$10^5, \dots, 10^2$	10^6	$10, \dots, 10^4$	10	10	$V \rightarrow \{10\}$	-
<i>Figure 10</i>							
Left, center	$10^3, \dots, 10^7$	$10^5, \dots, 10^9$	10^2	10	10	$V \rightarrow \{10\}$	-
Right	$10^2, \dots, 10^5$	$10^4, \dots, 10^7$	10^2	10	10	$V \rightarrow \{10\}$	-
<i>Figure 11</i>							
Left, right	$10^2, \dots, 10^5$	$10^4, \dots, 10^7$	10^2	10	10	$V \rightarrow \{10\}$	-
<i>Real-world graphs</i>							
Table 3	-	-	-	5	10	$V \rightarrow \{10\}$	-
Table 4	-	-	-	10	10	$V \rightarrow \{10\}$	-
Table 5	-	-	-	10	$5, 10$	$V \rightarrow \{\Delta\}$	-
Figure 12	-	-	-	10	$5, 20$	$V \rightarrow \{\Delta\}$	-
<i>Case study (Socio-patterns)</i>							
Figure 13	-	-	-	10	10	$V \rightarrow \{10\}$	5
Figure 14	-	-	-	10	10	$V \rightarrow \{10\}$	25
Figure 15	-	-	-	10	10	$V \rightarrow \{10\}$	5, 10, 25

Here n is the number of vertices, m is the number of edges, $d = \frac{m}{n}$ is the average degree, k is the length of path, Δ is maximum resting time, δ is the function mapping vertices to resting time, maximum timestamp $\tau = 100$ (fixed) and percentage of vertices considered as separators 'Sep (%)'. For power-law graphs, we use $\alpha = -1.0$ (fixed) and $w = 100$ (fixed). Recall that for a positive integer k we write $[k] = \{1, \dots, k\}$

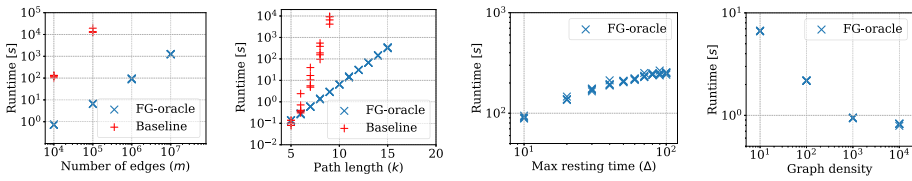


Fig. 9 Scalability in synthetic graphs. FG-oracle is fine-grained oracle

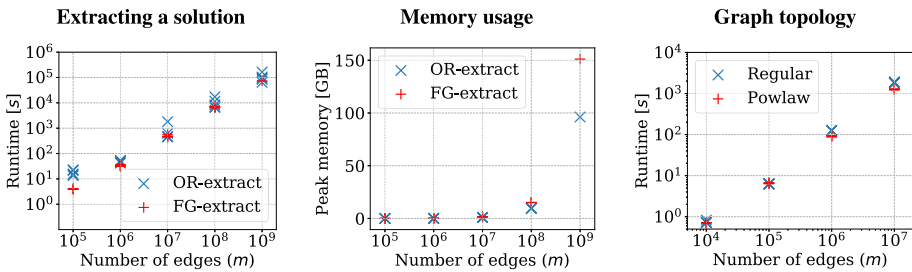


Fig. 10 Extracting a solution and topology invariance. OR-extract is solution extraction using decision oracle and FG-extract is solution extraction using fine-grained oracle

for sparse graphs there is not enough work to keep both the arithmetic and memory pipeline busy, simultaneously. This also presents us a challenge to design an efficient implementation for handling sparse graphs.

Extracting an optimal solution. Our second set of experiments compares the runtime of the algorithm for extracting an optimal restless path between a source and a destination using the decision oracle and the fine-grained oracle. Recall that the algorithm only returns a YES/NO answer for the existence of a solution and we need multiple queries to the oracle to obtain a solution. Using a decision oracle we require $\mathcal{O}(k \log n \log \tau)$ queries in expectation to obtain an optimum solution as compared to k queries using a fine-grained oracle (see Sect. 4.7). In Fig. 10, we compare the runtime (left) and the peak-memory usage (center) of the decision oracle and the fine-grained oracle for extracting an optimum solution for five independent random power-law graphs for each configuration of $n = 10^3, \dots, 10^7$ with fixed values of $d = 100, \tau = 100, \Delta = 10, \delta : V \rightarrow \{\Delta\}$ and $k = 10$. The source and the destination are chosen at random. We observe little variance in runtime for the fine-grained extraction as compared to oracle extraction. For large graphs with hundred million edges the fine-grained extraction is up to four times faster than the oracle extraction. Even though in theory we reduce the number of queries by a factor of $\log n \log \tau$, in practice we do not obtain a significant improvement in the empirical runtime. When extracting a solution using the decision oracle, recall that for each query of the oracle we recursively divide the graph into smaller subgraphs. For each smaller subgraph, we build the multilinear sieve and thereby reduce the empirical runtime of each query. In addition, while the expected number of queries is bounded by $\mathcal{O}(k \log n \log \tau)$ in worst case, this bound is not always met in practice. For a given instance, the number of queries required to extract a k -vertex subgraph varies depending on the source and the destination, resulting in high variance in the extraction time for extraction via self-reducibility. However, there is considerably less variance in the run time of the fine-grained extraction approach.

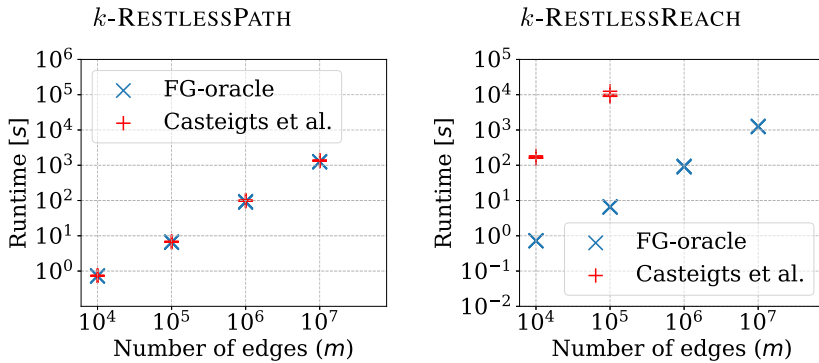


Fig. 11 Running time comparison of our fine-grained oracle and FPT(*k*)-algorithm of Casteigts et al. [13] for *k* - RESTLESSPATH and *k* - RESTLESSREACH problems. FG-oracle is fine-grained oracle

The experiments are executed on the compute-node configuration using all cores. We report the run time of extraction for *k*-RESTLESSMOTIF instances, but the run times are similar for *k*-RESTLESSPATH instances.

Graph topology. Here we study the effect of graph topology on run time of the algorithm. In Fig. 10 (right), we report the run time of the algorithm for five independent *d*-regular random graph instances for each configuration of $n = 10^2, \dots, 10^5$, $d = 100$, $\tau = 100$, $\Delta = 10$, $\delta : V \rightarrow \{\Delta\}$, and $k = 10$. Power-law graphs with $n = 10^2, \dots, 10^6$ with fixed values of $d = 100$, $k = 10$, $\tau = 100$, $\Delta = 10$, $\delta : V \rightarrow \{\Delta\}$, $\alpha = -1.0$, and $w = 100$. We observe no significant change in the runtime with the change in the graph topology. The experiments are executed on the workstation using all cores. We report the runtime for *k*-RESTLESSMOTIF-REACH instances, but the runtimes are similar for other problem instances.

Scalability of the FPT(*k*)-algorithm by Casteigts et al. [13]. We compare the runtime of our fine-grained oracle with the FPT(*k*)-algorithm by Casteigts et al. [13] for *k*-RESTLESSPATH and *k*-RESTLESSMOTIF problems. Figure 11 reports the runtime for both methods for five independent power-law graph instances for each configuration of $n = 10^2, \dots, 10^5$ with fixed values of $d = 100$, $k = 10$, $\tau = 100$, $\Delta = 10$, $\delta : V \rightarrow \{\Delta\}$, $\alpha = -1.0$, and $w = 100$. For RESTLESSPATH, we observe no significant difference in run time, however, for *k*-RESTLESSREACH the running time of Casteigts et al. [13] increased by a factor of *n*, consistent with theoretical expectations. The experiments are executed in the workstation configuration.

Experiments with real-world graphs. Our next set of experiments reports the run time of the algorithm for finding restless reachability and extracting a restless path in real-world datasets. The description of the datasets used for our experiments is available in Sect. 5.2.

In Tables 3 and 4, we report the runtime for finding restless reachability in real-worlds graphs with $k = 5$ and $k = 10$, respectively. For each dataset, we report the maximum runtime of five independent runs by choosing the source vertex $s \in V$ uniformly at random for each $k \in \{5, 10\}$, for fixed value of the maximum resting time $\Delta = 10$, and $\delta : V \rightarrow \{\Delta\}$, i.e., the resting time is constant for all the vertices except the source s , which has the maximum resting time $\delta(s) = \tau$. In Column 2, we report the runtime for solving *k*-RESTLESSREACH, while Column 3 reports the runtime for solving RESTLESSREACH by restricting the path length to $k - 1$. Note that here we solve *atm*-*k*-RESTLESSREACH where we find the set of vertices that are reachable from a given source via a restless path with length at most $k - 1$.

Table 3 Experiments with real-world datasets ($k = 5$)

Dataset	Reachability			Extraction ($k = 5$)			Memory(GB)
	$\bar{k} = 5$ (s)	$k \leq 5$ (s)	Ratio (s)	Fine-grained (s)	Decision (s)	Speedup (s)	
<i>Copenhagen</i>							
Calls	1.72	2.66	1.55	1.71	1.80	1.05	0.48
SMS	1.91	2.59	1.36	1.84	1.88	1.02	0.48
<i>Socio-patterns</i>							
LH10	8.18	13.06	1.60	8.24	8.27	1.00	3.28
InVS13	25.78	39.93	1.55	25.42	25.45	1.00	6.95
InVS15	39.21	59.38	1.51	42.30	38.36	0.91	13.18
SFHH	6.91	10.36	1.50	7.18	8.45	1.18	1.47
LyonSchool	22.39	32.56	1.45	22.58	27.33	1.21	1.70
Thiers13	51.76	73.39	1.42	51.46	55.27	1.07	5.38
<i>Koblenz</i>							
sqwikibooks	1.63	3.60	2.21	1.69	2.14	1.27	0.85
pswiktionary	6.42	10.67	1.66	6.84	7.28	1.06	3.70
sawikisource	9.96	17.21	1.73	13.60	16.63	1.22	6.58
knwiki	21.42	31.50	1.47	23.68	25.22	1.06	12.27
epinions	7.76	12.41	1.60	7.44	24.36	3.27	3.57
<i>Transport</i>							
Kuopio	0.49	0.78	1.57	0.51	0.86	1.70	0.10
Rennes	0.76	1.45	1.92	0.74	1.21	1.63	0.25
Grenoble	0.88	1.41	1.60	0.86	1.25	1.45	0.29
Venice	1.03	1.63	1.58	1.33	1.39	1.05	0.39
Belfast	0.82	2.31	2.81	0.89	1.31	1.47	0.31
Canberra	1.08	1.73	1.59	1.07	1.57	1.47	0.43
Turku	0.86	1.51	1.75	0.95	1.31	1.38	0.33
Luxembourg	0.86	1.19	1.38	0.97	1.20	1.24	0.24
Nantes	1.27	1.96	1.54	1.28	1.42	1.11	0.43
Toulouse	1.37	2.32	1.69	1.52	1.94	1.28	0.59
Palermo	1.25	1.86	1.49	1.30	1.57	1.21	0.40
Bordeaux	1.71	2.64	1.54	1.70	1.94	1.14	0.64
Antofagasta	0.70	0.87	1.23	0.68	1.58	2.34	0.11
Detroit	2.63	4.13	1.57	2.50	2.93	1.17	1.22
Winnipeg	2.33	3.33	1.43	2.31	2.69	1.17	0.94
Brisbane	3.92	5.62	1.43	3.77	4.12	1.09	1.76
Adelaide	3.22	4.72	1.47	3.04	3.58	1.18	1.36
Dublin	2.10	3.04	1.45	2.04	2.50	1.23	0.82
Lisbon	3.36	5.27	1.57	3.20	3.69	1.15	1.47
Prague	3.11	4.52	1.45	2.95	3.21	1.09	1.11
Helsinki	3.69	5.06	1.37	3.49	3.97	1.14	1.46

Table 3 continued

Dataset	Reachability			Extraction ($k = 5$)			Memory(GB)
	$k = 5$ (s)	$k \leq 5$ (s)	Ratio (s)	Fine-grained (s)	Decision (s)	Speedup (s)	
Athens	3.54	6.29	1.78	3.40	3.91	1.15	1.46
Berlin	2.95	3.93	1.33	3.31	3.69	1.11	1.01
Rome	4.59	6.01	1.31	4.44	4.99	1.13	1.70
Melbourne	8.24	12.61	1.53	8.56	8.18	0.96	4.00
Sydney	12.15	17.98	1.48	11.10	11.39	1.03	5.20
Paris	7.98	10.72	1.34	7.54	7.44	0.99	2.33

For a description of the columns, see the main text

Table 4 Experiments with real-world datasets ($k = 10$)

Dataset	Reachability			Extraction ($k = 10$)			Memory(GB)
	$k = 10$ (s)	$k \leq 10$ (s)	Ratio (s)	Fine-grained (s)	Decision (s)	Speedup (s)	
<i>Copenhagen</i>							
Calls	88.98	152.91	1.72	90.60	194.03	2.14	0.48
SMS	89.32	153.67	1.72	88.36	195.87	2.22	0.48
<i>Socio-patterns</i>							
LH10	411.13	719.49	1.75	417.43	532.08	1.27	3.28
InVS13	1292.91	2342.52	1.81	1309.36	1613.37	1.23	6.95
InVS15	1871.23	3414.07	1.82	1966.12	2155.92	1.10	13.18
SFHH	314.48	569.10	1.81	318.29	581.37	1.83	1.47
LyonSchool	1284.53	2294.68	1.79	1298.16	1849.85	1.42	1.70
Thiers13	2666.22	4687.66	1.76	2663.80	3233.05	1.21	5.38
<i>Koblenz</i>							
sqwikibooks	70.78	131.26	1.85	72.20	257.57	3.57	0.85
pswiktionary	267.32	478.48	1.79	283.14	909.79	3.21	3.71
sawikisource	464.38	856.73	1.84	495.80	1450.61	2.93	6.59
knwiki	909.25	1659.58	1.83	969.38	2995.04	3.09	12.28
epinions	286.11	522.78	1.83	375.55	3051.37	8.13	3.58
<i>Transport</i>							
Kuopio	25.08	43.45	1.73	25.42	116.30	4.58	0.10
Rennes	32.15	59.12	1.84	34.24	161.09	4.71	0.25
Grenoble	38.38	68.99	1.80	39.15	155.41	3.97	0.29
Venice	50.07	84.05	1.68	49.62	176.52	3.56	0.39
Belfast	36.29	69.19	1.91	39.32	172.56	4.39	0.31
Canberra	48.00	86.93	1.81	48.93	182.02	3.72	0.43
Turku	39.46	75.04	1.90	43.81	158.19	3.61	0.33
Luxembourg	37.15	65.73	1.77	39.68	140.88	3.55	0.24
Nantes	51.82	87.37	1.69	52.35	150.44	2.87	0.43
Toulouse	62.96	113.45	1.80	62.54	203.01	3.25	0.59

Table 4 continued

Dataset	Reachability			Extraction ($k = 10$)			Memory(GB)
	$k = 10$ (s)	$k \leq 10$ (s)	Ratio (s)	Fine-grained (s)	Decision (s)	Speedup (s)	
Palermo	47.20	93.09	1.97	52.74	198.76	3.77	0.40
Bordeaux	67.52	124.61	1.85	69.16	208.33	3.01	0.64
Antofagasta	23.72	41.53	1.75	25.26	112.66	4.46	0.11
Detroit	111.74	201.14	1.80	107.84	288.48	2.68	1.22
Winnipeg	88.13	162.17	1.84	93.29	231.43	2.48	0.94
Brisbane	147.66	268.12	1.82	153.42	320.82	2.09	1.76
Adelaide	122.11	221.79	1.82	124.29	312.47	2.51	1.37
Dublin	78.42	142.13	1.81	82.51	205.09	2.49	0.82
Lisbon	134.25	246.37	1.84	137.01	332.89	2.43	1.47
Prague	108.59	199.35	1.84	107.47	245.90	2.29	1.11
Helsinki	130.68	231.65	1.77	136.54	319.97	2.34	1.46
Athens	129.10	233.02	1.80	130.00	324.98	2.50	1.46
Berlin	102.28	179.79	1.76	105.72	233.19	2.21	1.01
Rome	159.08	287.87	1.81	164.57	368.35	2.24	1.70
Melbourne	323.32	590.11	1.83	337.50	631.20	1.87	4.00
Sydney	411.24	732.24	1.78	427.51	866.15	2.03	5.20
Paris	223.02	403.70	1.81	224.21	477.10	2.13	2.33

For description of the columns, please see the main text

Column 4 is the ratio of Column 3 and Column 2. Next, in Columns 5 and 6 we report the runtime of extracting a solution using a decision oracle and a fine-grained oracle, respectively. Column 7 is the ratio of Column 6 and Column 5. Finally, in Column 8, we report the peak memory usage of the fine-grained extraction. All run times are in seconds. The experiments are executed on compute-node configuration using all cores.

We can solve restless reachability in each real-world graph dataset in Table 7 in less than one minute by restricting the length of the restless path k to 5 and in less than two hours using at most 14 Gb of memory for $k = 10$. For instance, we can solve restless reachability by limiting k to 10 in a real-world graph dataset with more than 37 million directed edges and more than 19 thousand timestamps in less than one hour on a Haswell desktop using less than 6 Gb of memory. The reported runtimes are in seconds.

From the results (see Column 3), we see that solving *atm-k* - RESTLESSREACH takes less than twice the empirical running time than that of *k* - RESTLESSREACH, in most of the input graph instance. Extracting a solution using fine-grained extraction is effective for large graphs, given that k is not too large. For instance, we obtain an 8-time speedup in computation using fine-grained extraction as compared to oracle extraction in a graph with more than hundred thousand vertices and more than eight hundred thousand edges with $k = 10$.

Reachability. Our next set of experiments studies the restless reachability in socio-patterns dataset. In particular, we solve *atm-k* - RESTLESSREACH where we find the set of vertices reachable from a given source via a restless path of length at most $k - 1$. *Reachability* is the ratio of the number of vertices that are reachable with a restless path to the total number of vertices. In Fig. 12, we report the variance and mean of reachability as a time-series for five independent graph instances for each dataset with $\Delta = 5$ (top-row) and $\Delta = 20$ (bottom-

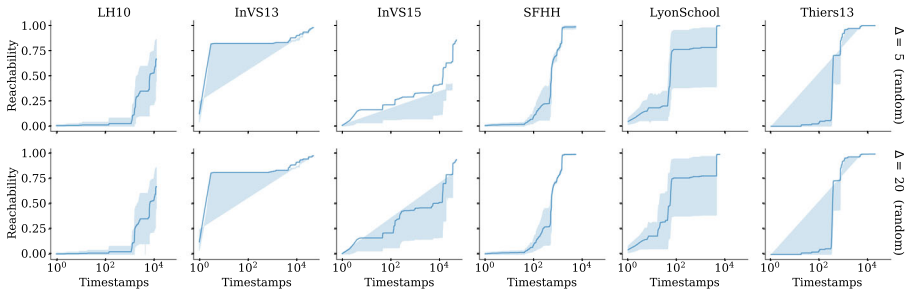


Fig. 12 Restless reachability in real-world datasets with $\Delta = 5$ (top-row) and $\Delta = 20$ (bottom-row) with $\delta : V \rightarrow [\Delta]$ assigned uniformly at random, and $k = 10$ (fixed)

Table 5 Computing restless reachability in socio-patterns dataset using the workstation configuration

Dataset	n	m	τ	$\Delta = 5$ (s)	$\Delta = 10$ (s)
LH10	73	300 252	12 960	156.8 s	158.7
InVS13	95	788 494	49 679	357.1 s	386.5
InVS15	219	2 566 388	49 679	657.2 s	733.2
SFHH	403	2 834 970	5 328	245.0 s	346.3
LyonSchool	242	13 188 984	5 887	853.6 s	1 341.2
Thiers13	328	37 226 078	19 022	2 083.1 s	3 399.5

We report run time for solving atm - k - RESTLESSREACH problem with $k = 10$ (fixed)

row). More precisely, for a given dataset we generate five graph instances by choosing a source vertex uniformly at random and assign the resting times uniformly at random in the range $[\Delta]$. We limit the length of the restless path to 9, i.e., $k = 10$ and solve atm - k - RESTLESSREACH.

We observe high variance in reachability among the independent source vertices for LH10, InVS15, LyonSchool, and Thiers13 datasets, and smaller variance for InVS13 and SFHH. We also see that in all datasets, except LH10, reachability approaches its maximum value 1, within the total number of timestamps available in each dataset, however, this happens at different times for each dataset, and with a different pace.

In Table 5, we report the run time for solving atm - k - RESTLESSREACH in socio-patterns datasets. The reported runtime is the maximum of five independent runs by choosing the source vertex $s \in V$ uniformly at random with $k = 10$ (fixed) and $\Delta = 5, 10, \delta : V \rightarrow \{\Delta\}$ (fixed). The resting time is constant for all the vertices except the source s , which has the maximum resting time $\delta(s) = \tau$. For instance, we can solve k - RESTLESSREACH by limiting $k = 10$ in a real-world graph dataset with more than 37 million directed edges and more than 19 thousand timestamps in less than one hour on workstation configuration using less than 6 Gb of memory. The reported runtimes are in seconds.

5.5 Case study: comparing immunization strategies

Our final set of experiments studies the change in restless reachability in the presence of a set of barrier vertices $S' \subseteq V$ called *separators*, which must not be included in the restless path. In an epidemic model, the separators can be viewed as a subset of the population, all immune

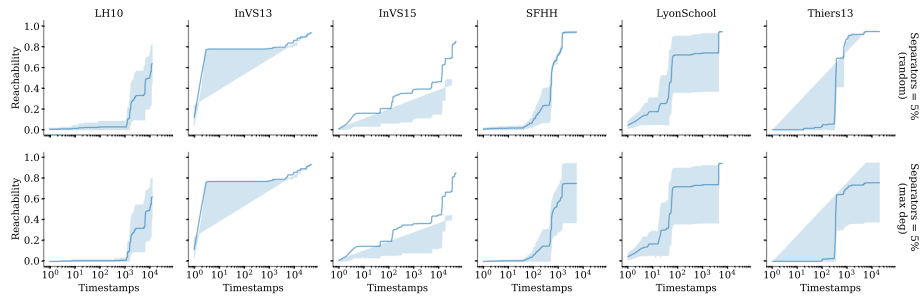


Fig. 13 Restless reachability in real-world datasets in the presence of 5% separators. $\Delta = 10, \delta : V \rightarrow [\Delta]$ assigned uniformly at random and $k = 10$ (fixed)

and/or vaccinated. It is known that finding a set of temporal separators with minimum size, which destroy all the restless paths between any two vertices, is \sum_2^P -hard [44][Theorem 7].

The experiments performed in this case study are to evaluate the effectiveness of our algorithm to answer queries for finding an effective immunization strategy in an epidemic model where the disease propagation is via a restless path. To demonstrate this, we use two heuristics for finding separators: (i) choose vertices at random and (ii) choose vertices with maximum temporal degree. Note that these immunization strategies are simple and without theoretical guarantees. Toward this end, we would like to investigate effective approximation or heuristic schemes to find temporal separators, which reduce the fraction of vertices reachable from a given source vertex via a restless path in future work.

Given an instance $(G = (V, E), s, k)$ of k -RESTLESSREACH and a set $S' \subseteq V$ of separators, we introduce a coloring function $c : V \setminus \{s' \cup s\} \rightarrow \{1\}, c(s) = 2$ and $c : S' \rightarrow \{3\}$ and $M = \{1^{\ell-1}, 2\}$ for $\ell \in [k]$. We query the FINEGRAINEDORACLE with instance $(G' = (V, E \setminus \{u, s, i\} \in E), c, \ell, M)$ for each $\ell \in \{2, \dots, k\}$. By assigning color $3 \notin M$ to the separators in S' , we make sure that none of the separators are part of the restless path agreeing the multiset of colors in M . Note that here we solve *atm*- k -RESTLESSREACH, in other words we find the set of vertices which are reachable from the source via a restless path of length at most $k - 1$.

In Fig. 13, we report the variance of reachability in real-world graphs by choosing 5% of the vertices uniformly a random (top-row) and 5% of the vertices with maximum degree (second-row) as separators. Figure 14 reports the same experiments with 5% replaced with 25%. For each dataset we generate five graph instances, choose source vertices at random and assign resting times uniformly at random in the range $[\Delta]$ for $\Delta = 10$. Note that we use the same instances in the reachability experiments described above. Reachability is the ratio of number of reachable vertices to the total number of vertices excluding separators. From the experimental results, we observe high variance in the reachability among independent source vertices and the rate of increase of the reachability with time varies depending on the source vertex. More importantly, even though we expect (empirically) that by choosing vertices with high degree as separators should decrease the reachability as compared to choosing the vertices at random, this is not true for all the datasets. We only observe this phenomenon in SFHH and Thiers13 datasets for choosing 5% of the vertices as separators and in InVS15 and LyonSchool datasets while choosing 25% of the vertices as separators.

In Fig. 15, we report the variance of reachability in real-world graphs by choosing 5%, 10%, and 25% of the vertices uniformly a random (top-row) and 5%, 10%, and 25% of the vertices with maximum degree (second-row) as separators. Again, for each dataset we gener-

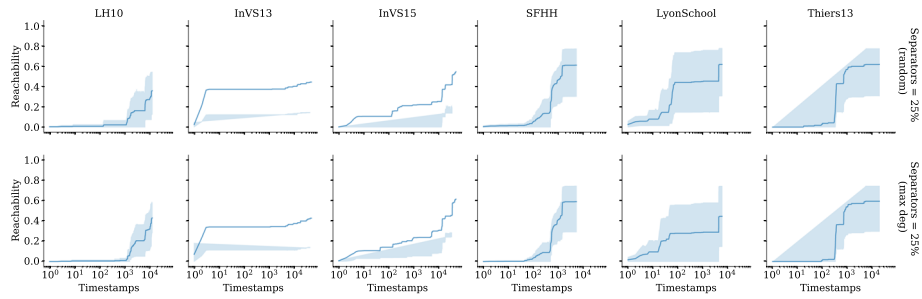


Fig. 14 Restless reachability in real-world datasets in the presence of 25% separators. $k = 10$, $\Delta = 10$, $\delta : V \rightarrow [\Delta]$ assigned uniformly at random

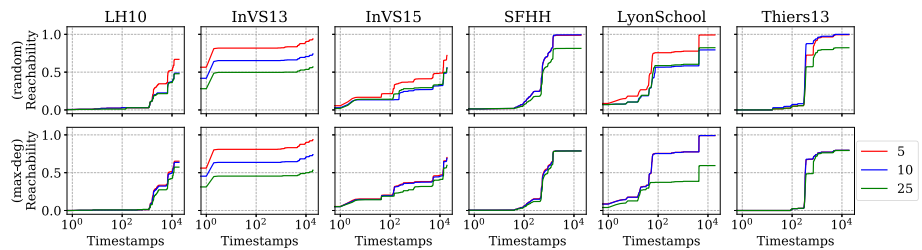


Fig. 15 Restless reachability in real-world datasets in the presence of separators. $k = 10$ (fixed), $\Delta = 10$, $\delta : V \rightarrow [\Delta]$ assigned uniformly at random

ate five graph instances, choose source vertices at random and assign resting time uniformly at random in the range $[\Delta]$ for $\Delta = 10$. Note that we use the same instances in the reachability experiments described above. We observe that reachability reduces with an increase in the number of separators, as expected. Again, even though it is expected (empirically) that choosing vertices with maximum degree to reduce the reachability significantly compared to choosing separators at random, surprisingly enough, this is not true for all datasets. For instance in LyonSchool dataset choosing 10% of the separators at random reduces the average reachability more than choosing vertices with maximum degree as separators. So the heuristic approach of choosing vertices with maximum degrees might not be effective across datasets. This presents us a interesting question of finding a small a set of separators in temporal graphs under resting time restrictions. Also note that the input graph datasets are small world graphs, meaning that the diameter of the underlying graph is small, so the vertices are highly connected (see Table 7).

6 Conclusions and future work

In this work, we studied a family of reachability problems in temporal and vertex-colored temporal graphs under waiting-time restrictions. We presented an algebraic algorithmic framework for solving restless reachability problems we proposed, running in $\mathcal{O}(2^k km \Delta)$ time and $\mathcal{O}(n \Delta)$ space. Further, we presented evidence that the algorithms for solving variants of the restless reachability problems involving colors presented in this work are optimal under certain complexity-theoretic assumptions. In addition, we engineered an open-source implementation of our algorithm and demonstrated its viability in experiments on graphs

with millions of temporal edges from real-world datasets. Finally, we applied the algorithm we developed in a case study for estimating the change in disease spreading with the presence of people with immunity. Our finding is that heuristic approaches such as selecting vertices with high degree as separators are not effective in all the graph datasets. Toward this end, we would like to investigate effective ways to choose a small set of separators under waiting-time restrictions to contain the spread of the disease in the network.

Future work We demonstrated that our algorithms scale to graphs with one billion edges for path lengths up to $k = 10$. Specifically, as our algorithms have running time exponential in k , i.e., $\mathcal{O}^*(2^k)$, the scalability with respect to k remains limited. Addressing this limitation by extending to larger k could be a direction of future work, potentially adapting the algorithm to vector-parallel architectures such as GPGPUs, as explored in prior work [32]. Further, we believe that our algorithms based on constrained multilinear sieving can be extended to solve other pattern-detection problems in temporal graphs, including finding temporal arborescences, connected temporal subgraphs, and temporal subgraphs with color constraints on the vertices. Additionally, we hypothesize that by extending the narrow-sieve construction proposed by Björklund et al. [5] to encode restless walks as a polynomial could break the $\mathcal{O}^*(2^k)$ barrier for k -RESTLESSPATH, potentially reducing the exponent.

Appendix

A Further motivation

Another application of restless reachability is to find signaling pathways in brain networks. Here we assume that brain regions are represented as vertices, and edges between the regions represent physical proximity, for example, two regions that are physically next to each other or having high signal correlation in an electro-encephalogram (EEG) are connected with an edge. Using functional magnetic resonance imaging (fMRI), we can record the *active* brain regions at different timestamps—the frequency of the scans is approximately one second [38, 56]. We introduce a timestamp on an edge if two regions with a static edge are active in consecutive fMRI scans. Finally, we introduce resting time for each region: any incoming signal can be forwarded to another region within the resting time duration. Given the source of signal origin we would like to find all the regions to which the signal was successfully transmitted. The problem can be abstracted as an instance of the RESTLESSREACH problem.

Also, consider a tour-recommendation scenario where a traveler is interested in visiting a set of places such as a historic museum, an art gallery, a café, or other. Each location has a maximum time-limit that the traveler is willing to spend. Given a start and an end location we would like to find a travel itinerary satisfying the constraints specified by the traveler. This problem can be modeled as an instance of the k -RESTLESSMOTIF problem, where each location can be considered as a vertex, the color associated with the vertex represents the type of a location, for example, museum, art gallery, etc., the temporal edges between the vertices represent the transportation links, and the resting time duration associated with each vertex represents the maximum time a traveler is willing to spend at a location.

B List of symbols

See Table 6

Table 6 List of symbols

Symbol	Description
<i>Graphs</i>	
$G = (V, E)$	Graph G with vertex set V and edge set E
n	Number of vertices
m	Number of edges
m_i	Number of edges at time stamp i
Δ	Maximum resting time
τ	Maximum timestamp
k	Length of path or walk
$\delta : V \rightarrow [\Delta]$	Restless function mapping vertices to resting time
$N_i(u)$	In-neighbors of vertex u at time stamp i
G^\downarrow	Underlying static graph of temporal graph G
$G^\uparrow(\delta)$	δ -expansion of temporal graph G
<i>Coloring</i>	
C	Set of colors
$c : V \rightarrow C$	Coloring function mapping vertices to colors
M	Multiset of colors
$\mu(s)$	Number of occurrences of color s in M
<i>Polynomials</i>	
\mathcal{P}	Polynomial encoding of temporal walks
χ	Polynomial encoding of restless walks
ζ	Evaluation polynomial of restless walks
x_u	x -variable encoding vertex u
$y_{uv,\ell,j}$	y -variable encoding edge (u, v, j) at position ℓ
\vec{x}	Vector of x variables
\vec{y}	Vector of y variables
\vec{z}	Vector of z variables

C Polynomial encoding temporal walks and temporal paths

Here we give brief overview of polynomial encoding of walks and paths in static as well as temporal graphs.

C.1 Monomial encoding of walks

Static graphs. We introduce a set of variables $\vec{x} = \{x_{v_1}, \dots, x_{v_n}\}$ for vertices in $V = \{v_1, \dots, v_n\}$ and a set of variables $\vec{y} = \{y_{uv,\ell} : (u, v) \in E, \ell \in [k]\}$ such that $y_{uv,\ell}$ corresponds to an edge $(u, v) \in E$ that appears at position ℓ in a walk. Using these variables a walk $W = v_1 e_1 v_2 \dots e_{k-1} v_k$ can be encoded by the monomial

$$x_{v_1} y_{v_1 v_2, 1} x_{v_2} \dots y_{v_{k-1} v_k, k-1} x_{v_k}.$$

The monomial encoding of walks in static graphs is illustrated in Fig. 16. It is easy to see that the encoding is multilinear, i.e., no variable in a monomial is repeated, if and only if the

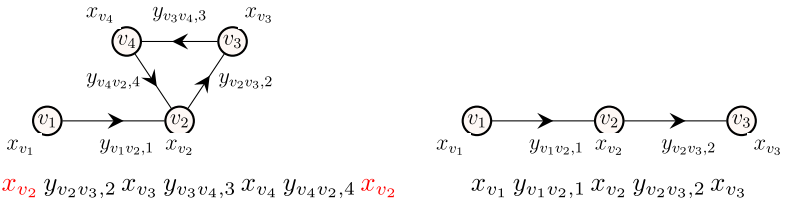


Fig. 16 A monomial encoding of a walk (left) and a path (right) in static graphs. Arrows indicate the direction of the walk. Variables highlighted in red are repeated if and only if a walk is not a path

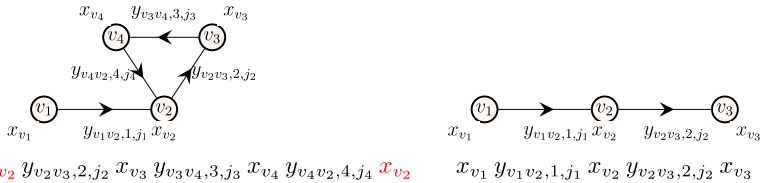


Fig. 17 A monomial encoding of a temporal walk (left) and a temporal path (right). Arrows indicate the direction of the walk. Variables highlighted in red are repeated if and only if a temporal walk is not a temporal path

corresponding walk is a path. To encode a walk of length $k - 1$, we need k variables of \vec{x} and $k - 1$ variables of \vec{y} , for a total of $2k - 1$ variables.

Temporal graphs. We introduce a set of variables $\vec{x} = \{x_{v_1}, \dots, x_{v_n}\}$ for vertices $V = \{v_1, \dots, v_n\}$ and a set of variables $\vec{y} = \{y_{uv,\ell,j} : (u, v, j) \in E, \ell \in [k]\}$ such that $y_{uv,\ell,j}$ corresponds to an edge $(u, v, j) \in E$ that appears at position ℓ in a walk. Using these variables, a temporal walk $W = v_1 e_1 v_2 \dots e_{k-1} v_k$ can be encoded using the monomial

$$x_{v_1} y_{v_1 v_2, 1, j_1} x_{v_2} y_{v_2 v_3, 2, j_2} \dots y_{v_{k-1} v_k, k-1, j_{k-1}} x_{v_k}.$$

The polynomial encoding of temporal walks is illustrated in Fig. 17. A monomial encoding of a temporal walk of length $k - 1$ has $2k - 1$ variables.

Again, observe that the encoding is multilinear if and only if the corresponding temporal walk is a temporal path. In fact, it can be shown that a temporal walk is a temporal path if and only if the corresponding monomial encoding is multilinear [53, Lemma 2]. Likewise, the problem of deciding the existence of temporal path of length $k - 1$ is equivalent to deciding the existence of a multilinear monomial of degree $2k - 1$ [53, Lemma 3]. In other words, if we encode all the walks of length $k - 1$ in a temporal graph using a polynomial where each monomial encodes a walk, then the problem of detecting the existence of a temporal path of length $k - 1$ is equivalent to the problem of detecting existence of a multilinear monomial of degree $2k - 1$ in the encoded polynomial.

C.2 Generating temporal walks

To generate a polynomial encoding of temporal walks, we turn to a dynamic-programming recursion. That is, we make use of polynomial encoding of temporal walks of length $\ell - 2$ to generate the encoding of temporal walks of length $\ell - 1$ recursively.

Consider the example illustrated in Fig. 18, where the vertex u has incoming edges from vertices $\{v_1, v_2\}$ at timestamp i , i.e., $N_i(u) = \{v_1, v_2\}$. Let $\mathcal{P}_{u,\ell,i}(\vec{x}, \vec{y})$ denote the polynomial encoding of all temporal walks of length $\ell - 1$ ending at vertex u at latest time i . Following our

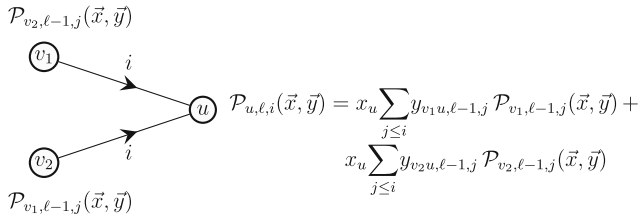


Fig. 18 Polynomial encoding of temporal walks

notation, $\mathcal{P}_{v_1, \ell-1, i}(\vec{x}, \vec{y}), \mathcal{P}_{v_2, \ell-1, i}(\vec{x}, \vec{y})$ represent the polynomial encoding of walks ending at vertices v_1, v_2 , respectively, such that all walks have length $\ell - 1$ and end at latest time i . From the definition of a temporal walk, it is clear that we can walk from vertices v_1, v_2 to vertex u at time i if we have reached v_1 and/or v_2 at latest time i . The polynomial encoding of walks of length $\ell - 1$ ending at vertex u at latest time i can be written as

$$\mathcal{P}_{u, \ell, i}(\vec{x}, \vec{y}) = x_u \sum_{j \leq i} y_{v_1 u, \ell-1, j} \mathcal{P}_{v_1, \ell-1, j}(\vec{x}, \vec{y}) + x_u \sum_{j \leq i} y_{v_2 u, \ell-1, j} \mathcal{P}_{v_2, \ell-1, j}(\vec{x}, \vec{y}).$$

By generalizing the above intuition, a generating polynomial for temporal walks can be written as:

$$\begin{aligned} \mathcal{P}_{u, 1, i}(\vec{x}, \vec{y}) &= x_u, \text{ for each } u \in V \text{ and } i \in [\tau], \text{ and} \\ \mathcal{P}_{u, \ell, i}(\vec{x}, \vec{y}) &= x_u \sum_{v \in N_i(u)} \sum_{j \leq i} y_{vu, \ell-1, j} \mathcal{P}_{v, \ell-1, j}(\vec{x}, \vec{y}) \end{aligned} \tag{5}$$

for each $u \in V, \ell \in \{2, \dots, k\}$ and $i \in [\tau]$.

D Dataset statistics

For each real-world dataset, we report the number of vertices n , the maximum timestamp τ , the number of temporal edges m , and the average temporal degree $d_{\text{avg}} = \frac{m}{n\tau}$. Furthermore, for each dataset we construct a static undirected graph or an *underlying graph* by considering the edges without the timestamp information ignoring multiple edges (if any) between the same set of vertices. For a temporal graph G , we denote the corresponding underlying (static) graph as G^\downarrow . For the underlying graph, we report the number of edges m^\downarrow , the average degree $d_{\text{avg}}^\downarrow$, the length of the diameter $D_{\text{max}}^\downarrow$, the feedback edge number (FEN), and the feedback vertex number (FVN). The dataset statistics are reported in Table 7.

The *average degree* $d_{\text{avg}}^\downarrow = \frac{m^\downarrow}{n}$ is the ratio of number of edges and the number of vertices in the static graph. The *diameter* of a static graph is the length of a longest shortest path between any two vertices and it can be computed in time $\mathcal{O}(nm^\downarrow)$, where n is number of vertices and m^\downarrow is number of edges, however, the runtime can be further reduced to $\mathcal{O}(m^\downarrow)$ in practice [17]. The diameter of the underlying graph is neither a lower bound nor an upper bound on the maximum length of the restless path. We report this value only to show that usually the path length parameter is orders of magnitude less than the parameters FVN and FEN in most real-world graph datasets.

The *feedback edge number* of a static graph $G = (V, E)$ is the minimum number of edges that need to be removed from G in order to make G acyclic. Formally, a set $F \subseteq E$ is a

Table 7 Dataset statistics

Dataset	Temporal graph			Static undirected graph			FEN	FVN	
	n	m	τ	d_{avg}	$m \downarrow$	$d_{avg} \downarrow$			$D_{max} \downarrow$
<i>Copenhagen</i>									
Calls	536	3600	4028	0.00	621	1.16	22	248	50
SMS	568	24,333	4032	0.01	697	1.23	20	233	58
<i>Socio-patterns</i>									
LH10	73	300,252	12,960	0.31	1381	18.92	3	1309	50
InVS13	95	788,494	49,679	0.16	3915	41.21	2	3821	85
InVS15	219	2,566,388	49,679	0.24	16,725	76.37	3	16,508	198
SFHH	403	2,834,970	5328	1.32	73,557	182.52	2	73,155	384
LyonSchool	242	13,188,984	5887	9.26	26,594	109.89	2	26,353	235
Thiers13	328	37,226,078	19,022	5.96	43,496	132.61	2	43,169	315
<i>Koblenz</i>									
sqwikibooks	6607	21,709	906	0.00	7731	1.17	7	6693	22
pswiktionary	26,595	66,112	982	0.00	53,167	2.00	10	39,591	30
sawikisource	48,960	106,991	948	0.00	72,608	1.48	4	60,588	83
knwiki	78,142	714,594	1107	0.00	306,077	3.92	9	254,671	1965
epinions	131,828	841,372	189	0.01	711,783	5.40	16	652,338	-
<i>Transport</i>									
Kuopio	549	32,122	1232	0.05	699	1.27	29	219	43
Rennes	1407	109,075	1223	0.06	1670	1.19	57	329	95
Grenoble	1547	114,492	1314	0.06	1679	1.09	82	489	50
Venice	1874	118,519	1474	0.04	2647	1.41	61	875	222
Belfast	1917	122,693	1132	0.06	2180	1.14	64	291	95
Camberra	2764	124,305	1095	0.04	3206	1.16	51	475	135

Table 7 continued

Dataset	Temporal graph			Static undirected graph				FVN	
	n	m	τ	d_{avg}	$m \downarrow$	$d_{avg} \downarrow$	$D_{max} \downarrow$		FEN
Turku	1850	133,512	1260	0.06	2335	1.26	67	665	157
Luxembourg	1367	186,752	1211	0.11	1903	1.39	34	687	160
Nantes	2353	196,421	1280	0.07	2743	1.17	82	524	145
Toulouse	3329	224,516	1233	0.05	3734	1.12	111	542	147
Palermo	2176	226,215	1270	0.08	2559	1.18	90	384	123
Bordeaux	3435	236,595	1307	0.05	4026	1.17	98	668	211
Antofagasta	650	293,921	1097	0.41	963	1.48	51	327	94
Detroit	5683	214,863	1510	0.03	5946	1.05	206	647	100
Winnipeg	5079	333,882	1296	0.05	5846	1.15	84	818	259
Brisbane	9645	392,805	1283	0.03	11,681	1.21	164	2306	607
Adelaide	7548	404,300	1270	0.04	9234	1.22	89	1992	524
Dublin	4571	407,240	1256	0.07	5537	1.21	81	1559	319
Lisbon	7073	526,179	1457	0.05	8817	1.25	101	2158	-
Prague	5147	670,423	1510	0.09	6714	1.30	67	2446	470
Helsinki	6986	686,457	1465	0.07	9022	1.29	74	2401	655
Athens	6768	724,851	1506	0.07	7978	1.18	131	1375	399
Berlin	4601	1,048,218	1520	0.15	6600	1.43	46	2245	575
Rome	7869	1,051,211	1506	0.09	10,068	1.28	100	2519	686
Melbourne	19,493	1,098,227	1441	0.04	21,434	1.10	254	2531	665
Sydney	24,063	1,265,135	1519	0.03	28,695	1.19	115	5176	1511
Paris	11,950	1,823,872	1359	0.11	13,726	1.15	159	3736	613

Here, n is the number of vertices, m is the number of temporal edges, τ is the number of timestamps, d_{avg} is the average degree of the temporal graph, and $m \downarrow$ is number of edges, $d_{avg} \downarrow$ is the average degree, $D_{max} \downarrow$ is the diameter, FEN is the feedback edge number, FVN is the feedback vertex number of the underlying static graph. See main text for a precise definition of these parameters

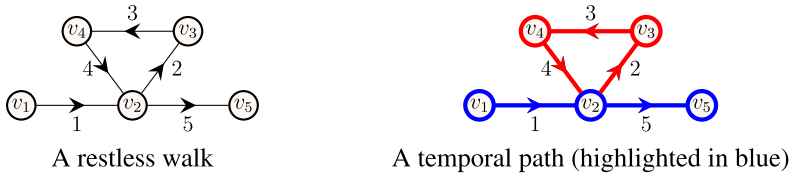


Fig. 19 An illustration of infeasibility to transform a restless walk to a restless path. On the left, a restless (temporal) walk $v_1 e_{v_1,v_2,1} v_2 e_{v_2,v_3,2} v_3 e_{v_3,v_4,3} v_4 e_{v_4,v_2,4} v_2 e_{v_2,v_5,5} v_5$ from vertex v_1 to v_5 , where resting time of vertices is $\delta(v_1) = \dots = \delta(v_5) = 2$. On the right, a temporal path $v_1 e_{v_1,v_2,1} v_2 e_{v_2,v_5,5} v_5$ from v_1 to v_5 (highlighted in blue) obtained by removing the loop (highlighted in red). Observe that the obtained temporal path is not a restless path since the difference in edge timestamps entering and leaving v_2 is $4 > \delta(v_2) = 2$. In conclusion, even though there is a restless walk from v_1 to v_5 , there exists no restless path from v_1 to v_5

feedback edge set if $(V, E \setminus F)$ does not contain a cycle. In other words, the feedback edge number of G is the size of a minimum feedback edge set for G . The feedback edge number can be computed in polynomial-time by computing a maximum spanning forest of G and by taking its edge-complement.

The *feedback vertex number* of a static graph $G = (V, E)$ is the minimum number of vertices that need to be removed from G in order to make G acyclic. Formally, a set $F \subseteq V$ is a *feedback vertex set* if $(V \setminus F, E \setminus \{(u, v) : u \in F \text{ or } v \in F\})$ does not contain a cycle. The feedback vertex number of G is the size of minimum feedback vertex set of G . In contrast to the feedback edge number, computing the feedback vertex number is **NP-hard** (see e.g., [31]). For computing the feedback vertex number, we use the solvers from the work of Iwata et al. [28, 29].

E Baselines

In this section, we discuss the choice of baselines considered for comparison.

Random restless walks. Algorithmic approaches based on random walks can be used to estimate reachability in temporal graphs. The approach works by performing a random walk in a temporal graph and update reachability using the transitive closure property by respecting time constraints. More precisely, let u, v and w be distinct vertices. Now, if there exists a temporal walk from u to v ending at time i and a temporal walk starting at time j from v to w such that $i \leq j$, then there is a temporal walk from u to w . Most reachability methods using temporal walks can be extended to estimate reachability with temporal paths, since a temporal walk can be transformed into a temporal path by removing loops in the walk. However, a similar approach of transforming reachability using restless walks to restless paths is not straightforward, since a restless walk cannot be transformed into a restless path by simply removing loops, as it may not satisfy waiting-time constraints (see Fig. 19).

Temporal graph expansion. A time-expansion of a temporal graph to a static directed graph has been applied to reduce temporal reachability to reachability questions in static directed graphs [13, 42, 52, 60, 64]. Again, the approach can be used to solve reachability problems when vertices need to be connected via a restless walk; however, it fails to solve reachability problems when vertices must be connected via a restless path. Recall that we described a transformation of a temporal graph G to a static directed graph $G^\uparrow(\delta)$ which respects waiting-time restrictions called δ -expansion in Sect. 4.8. For an illustration, see Fig. 8.

From the construction of δ -expansion, it is easy to see that there exists a restless walk from vertex v_i to v_j in G if and only if there exist a directed path from vertex v_i^ℓ to $v_j^{\ell'}$ in $G^\uparrow(\delta)$ for some $\ell, \ell' \in [\tau]$ such that $\ell \leq \ell'$. Using δ -expansion, we can use existing algorithmic approaches for solving reachability problems in static directed graphs to solve reachability problems where vertices must be connected via a restless walk. Observe that even though there exists a directed path from v_1^1 to v_5^5 in $G^\uparrow(\delta)$, there is no restless path from v_1 to v_5 in G , so it is not straightforward to employ this approach to solve reachability problems when vertices must be connected via a restless path. Even though we presented a deterministic algorithm to solve RESTLESSPATH via a transformation to RAINBOWPATH, the algorithm is not practical for graphs with large number of vertices.

Index construction. In a 2-hop labeling of static graphs, each vertex $v \in V$ is assigned a label-set pair $(L_{\text{in}}(v), L_{\text{out}}(v))$ such that v is reachable from each $u \in L_{\text{in}}(v) \subseteq V$ and each $w \in L_{\text{out}}(v) \subseteq V$ is reachable from v . From the transitive closure property, it follows that a vertex v is reachable from a vertex u if and only if $L_{\text{out}}(u) \cap L_{\text{in}}(v) \neq \emptyset$. Computing a 2-hop labeling of minimum size is **NP-hard** [14, Theorem 4.1], while approximable within a logarithmic factor [14, Theorem 4.2].

The approach is extended to solve temporal reachability using a transformation to a directed acyclic graph (DAG) [61, Section 3]. Here, the authors try to answer reachability questions in temporal graphs constrained by time intervals. More specifically, given two vertices $u, v \in V$ and timestamps $t_1, t_2 \in [\tau]$, the goal is to decide whether v is reachable from u via a temporal path (or a temporal walk) such that the timestamps of the edges in the path are in range $\{t_1, t_1 + 1, \dots, t_1 + t_2\}$. A key difference in our reachability model is that the transition time of an edge can be zero which makes the δ -expansion described in Fig. 8 a static directed graph but not a DAG. Additionally, we established that even though two vertices are reachable via a directed path in the δ -expansion, it does not imply that there exists a restless path connecting the vertices in the temporal graph.

We would like to note that RESTLESSPATH is **NP-complete** for all integers $\Delta \geq 1$ and $\tau \geq \Delta + 2$, even if there exists at most one temporal edge between any two vertices [13, Theorem 5]. So it is highly unlikely that there exists a straightforward extension of 2-hop cover to solve RESTLESSPATH even if we constrain time intervals.

Parameterized (exact) algorithms. Casteigts et al. [13] presented algorithms for solving RESTLESSPATH parameterized by the feedback edge number of the underlying graph and the timed feedback vertex number (TFVN) of the temporal graph. Additionally, they showed that the TFVN of the temporal graph is lower bounded by the FVN of the underlying graph. For a temporal graph $G = (V, E)$ with maximum timestamp τ , the *timed feedback vertex set* of G is a set $F \subseteq V \times [\tau]$ of vertex appearances such that $(G - F)^\downarrow$ is acyclic, where $G - F = (V, E')$ such that $E' = E \setminus \{(u, v, i) \in E : (v, i) \in F \text{ or } (v, i) \in F\}$ and $(G - F)^\downarrow$ is the underlying graph of $(G - F)$. The *timed feedback vertex number* of a temporal graph G is the minimum cardinality of a timed feedback vertex set of G .

From the statistics in Table 7, it is clear that the values of FVN and FEN are orders of magnitude greater than the diameter of the graph, for most datasets. Most importantly, it appears difficult to change the value of the parameters FEN and FVN, whereas the length of the restless path k to be found can be readily varied. Indeed, as we will see, such an approach leads to a solution scaling to graphs with millions of temporal edges, provided that the parameter k remains small enough. Also note that the deterministic algorithm presented in Sect. 4.8 is not practical for graphs with large value of n considered for experiments.

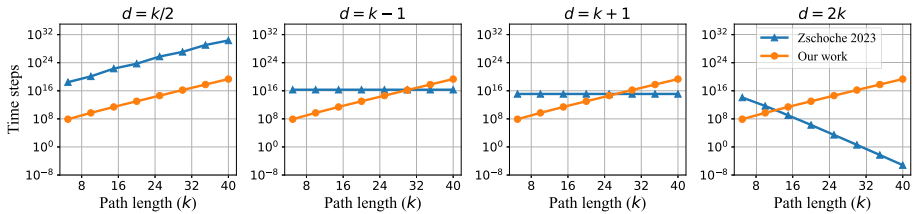


Fig. 20 A comparison of number of time steps of our algorithm and Zschoche [63] for solving k - RESTLESS-PATH

We implemented the $FPT(k)$ -algorithm by Casteigts et al. [13], which uses a different polynomial construction, with running time $\mathcal{O}(2^k(kn + km\Delta))$ and space complexity $\mathcal{O}(kn\tau)$.¹⁰ While the run time for a single execution of the decision oracle to decide the existence of a k - RESTLESSPATH between a source and a target is comparable to our algorithm, however, solving k - RESTLESSREACH requires $\mathcal{O}(n)$ queries to identify all reachable vertices, increasing the running time by a factor of n .

In a recent work, Zschoche [63] gave a randomized $\mathcal{O}(4^{k-d}(k-d)^2m^3\Delta)$ -time algorithm to find a shortest restless path (i.e., one minimizing the length of the path) between a source s and destination z , where d is the length of a shortest temporal path from s to z . For their algorithm to outperform ours, the values of k must be large. In Fig. 20, we compare the time complexity of our algorithm—i.e., $\mathcal{O}(2^kkm\Delta)$ —with that of Zschoche [63] for various configurations: $d = \frac{k}{2}$ (left), $d = k - 1$ (center-left), $d = k + 1$ (center-right), $d = 2k$ (right), for $k = \{5, 10, \dots, 40\}$, $n = 10^4$, $m = 10^6$, $\Delta = 5$ and $\tau = 100$. For $d = 2k$, we observe that the run time of both methods is similar for $k = 12$ and Zschoche’s algorithm outperforming ours for $k > 12$. However, our implementation struggles to scale for $k > 10$, even for moderately sized graphs with $m = 10^6$, making the practical runtime comparisons infeasible for large k .¹¹

Author Contributions The author contributions are as follows: We refer to authors Suhas Thejaswi, Juho Lauri, and Aristides Gionis as Authors 1, 2, and 3, respectively. Author 1 is responsible for formalizing the problem and design of algorithms, in discussions with Authors 2 and 3. Author 1 wrote the preliminary technical write-up, implemented and experimented with the algorithms. Authors 2 and 3 are also responsible for formalizing the proofs and writing the final draft of this article.

Funding Open Access funding enabled and organized by Projekt DEAL. Suhas Thejaswi acknowledges support from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (945719), European Union’s SoBigData++ Transnational Access Scholarship, and Nokia Foundation Scholarship (20220290). Aristides Gionis is supported by the ERC Advanced Grant REBOUND (834862), the EC H2020 RIA project SoBigData++ (871042), and the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation.

Data and code availability We use real-world data from several sources: the co-presence dataset from socio-patterns [23], Koblenz network collections [40], Copenhagen study network [49], and public-transport networks [39]. These datasets are publicly available and anonymized ensuring that no individual can be identified. Our implementations are available as open source [55] and are provided under a modest license agreement.

¹⁰ The exact polynomial factors and space complexity are not detailed in their paper. Instead, the reported complexities are based on our calculations.

¹¹ The results in Fig. 20 represent the number of time steps calculated based on their respective theoretical time complexities. Although practical execution times may vary, the runtime scaling of both algorithms is expected to follow a similar pattern in practice.

Declarations

Conflict of interest Authors declare no conflict of interest.

Ethical approval Our work is theoretical in nature, but we conducted experimental evaluation and case studies to 33 demonstrate practical applicability of our work. Further, our work does not involve any human subject study and/or 34 crowd-sourcing, as such approval from ethics committee and/or an institutional review board is not mandated for this 35 research.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Alon N, Dao P, Hajirasouliha I, Hormozdiari F, Cenk Sahinalp S (2008) Biomolecular network motif counting and discovery by color coding. *Bioinformatics* 24(13):241–249
2. Anand R, Chatterjee S (2017) Tracking disease progression by searching paths in a temporal network of biological processes. *PLoS One* 12(4):e0176172
3. Bakshy E, Rosenn I, Marlow C, Adamic L (2012) The role of social networks in information diffusion. In: Proceedings of the international conference on world wide web. ACM, pp 519–528
4. Benson E, Mohammed A, Gardell J, Masich S, Czeizler E, Orponen P, Höglberg B (2015) DNA rendering of polyhedral meshes at the nanoscale. *Nature* 523(7561):441–444
5. Björklund A, Husfeldt T, Kaski P, Koivisto M (2017) Narrow sieves for parameterized paths and packings. *J Comput Syst Sci* 87(2017):119–139
6. Björklund A, Kaski P, Kowalik Ł (2013) Probably optimal graph motifs. In: Proceedings of the international symposium on theoretical aspects of computer science. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, pp 20–31
7. Björklund A, Kaski P, Kowalik Ł (2014a) Determinant sums for undirected hamiltonicity. *SIAM J Comput* 43(1):280–299
8. Björklund A, Kaski P, Kowalik Ł (2014) Fast witness extraction using a decision oracle. In: Proceedings of the European symposium on algorithms, Springer, pp 149–160
9. Björklund A, Kaski P, Kowalik Ł (2016) Constrained multilinear detection and generalized graph motifs. *Algorithmica* 74(2):947–967
10. Björklund A, Kaski P, Kowalik Ł, Lauri J (2015) Engineering motif search for large graphs. In: Proceedings of the workshop on algorithm engineering and experiments. SIAM, pp 104–118
11. Bollobás B (2001) *Random graphs* (2nd ed.). Cambridge UP. ISBN0-521-80920-7; 0-521-79722-5
12. Bullmore E, Sporns O (2009) Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci* 10(3):186–198
13. Casteigts A, Himmel A-S, Molter H, Zschoche P (2021) Finding temporal paths under waiting time constraints. *Algorithmica* 83(9):2754–2802
14. Cohen E, Halperin E, Kaplan H, Zwick U (2003) Reachability and distance queries via 2-hop labels. *SIAM J Comput* 32(5):1338–1355
15. Cohen J, Italiano GF, Manoussakis Y, Nguyen KT, Pham HP (2021) Tropical paths in vertex-colored graphs. *J Comb Optim* 42(3):476–498
16. Cooke KL, Halsey E (1966) The shortest route through a network with time-dependent internodal transit times. *J Math Anal Appl* 14(3):493–498
17. Crescenzi P, Grossi R, Habib M, LANZI L, Marino A (2013) On computing the diameter of real-world undirected graphs. *Theor Comput Sci* 514:84–95
18. Cygan M, Dell H, Lokshtanov D, Marx D, Nederlof J, Okamoto Y, Paturi R, Saurabh S, Wahlström M (2016) On problems as hard as CNF-SAT. *ACM Trans Algorithms* 12(3):1–24
19. Cygan M, Fomin FV, Kowalik Ł, Lokshtanov D, Marx D, Pilipczuk M, Pilipczuk M, Saurabh S (2015) *Parameterized algorithms*, vol 4. Springer, Berlin

20. de Pasquale F, Penna SD, Snyder AZ, Lewis C, Mantini D, Marzetti L, Belardinelli P, Ciancetta L, Pizzella V, Romani GL, Corbetta M (2010) Temporal dynamics of spontaneous MEG activity in brain networks. *Proc Natl Acad Sci* 107(13):6040–6045
21. Dean BC (2004) Algorithms for minimum-cost paths in time-dependent networks with waiting policies. *Networks* 44(1):41–46
22. Dondi R, Hosseinzadeh MM (2023) Colorful path detection in vertex-colored temporal. *Netw Sci* 11(4):615–631
23. Génois M, Barrat A (2018) Can co-location be used as a proxy for face-to-face contacts? *EPJ Data Sci* 7(1):11
24. Glerean E, Pan RK, Salmi J, Kujala R, Lahnakoski JM, Roine U, Nummenmaa L, Leppämäki S, Wendt TN, Tani P, Saramaki J, Sams M, Jääskeläinen I (2016) Reorganization of functionally connected brain subnetworks in high-functioning autism. *Hum Brain Mapp* 37(3):1066–1079
25. Himmel A-S, Bentert M, Nichterlein A, Niedermeier R (2019) Efficient computation of optimal temporal walks under waiting-time constraints. In: *Proceedings of the international conference on complex networks and their applications*. Springer, pp 494–506
26. Holme P (2005) Network reachability of real-world contact sequences. *Phys Rev E* 71(4):046119
27. Holme P, Saramäki J (2012) Temporal networks. *Phys Rep* 519(3):97–125
28. Iwata Y, Wahlstrom M, Yoshida Y (2016) Half-integrality, LP-branching, and FPT algorithms. *SIAM J Comput* 45(4):1377–1411
29. Iwata Y, Wahlstrom M, Yoshida Y (2018) <https://github.com/wata-orz/fvs>
30. Jackson Matthew O, Yariv L (2006) Diffusion on social networks. *Economie publique/Public economics* 16
31. Karp Richard M (1972) Reducibility among combinatorial problems. In: *Proceedings of the symposium on the complexity of computer computations*. Plenum Press, pp 85–103
32. Kaski P, Lauri J, Thejaswi S (2018) Engineering motif search for large motifs. In: *Proceedings of the international symposium on experimental algorithms*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, vol 28, pp 1–28
33. Kermack WO, McKendrick AG (1927) A contribution to the mathematical theory of epidemics. *R Soc Lond, Series A* 115(772):700–721
34. Koutis I (2008) Faster algebraic algorithms for path and packing problems. In: *Proceedings of the international colloquium on automata, languages and programming (LNCS)*. Springer, pp 575–586
35. Koutis I, Williams R (2009) Limits and applications of group algebras for parameterized problems. In: *Proceedings of the international colloquium on automata, languages and programming (LNCS)*. Springer, pp 653–664
36. Kowalik Ł, Lauri J (2016) On finding rainbow and colorful paths. *Theor Comput Sci* 628(2016):110–114
37. Krauthgamer R, Trabelsi O (2019) The set cover conjecture and subgraph isomorphism with a tree pattern. In: *Proceedings of the international symposium on theoretical aspects of computer science*. Dagstuhl-Leibniz-Zentrum fuer Informatik, vol 45, pp 1–45
38. Kujala R (2018) From the brain to public transport: applications of network science. Ph. D. Dissertation. Aalto University
39. Kujala R, Weckström C, Darst RK, Mladenović MN, Saramäki J (2018) A collection of public transport network data sets for 25 cities. *Sci Data* 5(2018):180089
40. Kunegis J (2013) KONECT: the Koblenz network collection. In: *Proceedings of the International world wide web conference*. ACM, pp 1343–1350
41. Lin S-J, Al-Naffouri TY, Han YS, Chung WH (2016) Novel polynomial basis with fast Fourier transform and its application to Reed-Solomon erasure codes. *IEEE Trans Inf Theory* 62:11
42. Mertzios GB, Michail O, Spirakis PG (2019) Temporal network optimization subject to connectivity constraints. *Algorithmica* 81(4):1416–1449
43. Michail O (2016) An introduction to temporal graphs: an algorithmic perspective. *Internet Math* 12(4):239–280
44. Molter H (2022) The complexity of finding temporal separators under waiting time constraints. *Inform Process Lett* 175(2022):106229
45. Nightingale EB, Douceur JR, Orgovan V (2011) Cycles, cells and platters: an empirical analysis of hardware failures on a million consumer PCs. In: *Proceedings of the international conference on computer systems (EuroSys '11)*. ACM, pp 343–356
46. Pastor-Satorras R, Castellano C, Van Mieghem P, Vespignani A (2015) Epidemic processes in complex networks. *Rev Modern Phys* 87(3):925
47. Perozzi B, Akoglu L, Iglesias Sánchez P, Müller E (2014) Focused clustering and outlier detection in large attributed graphs. In: *Proceedings of the international conference on Knowledge discovery and data mining*. ACM, pp 1346–1355

48. Rozenstein P, Gionis A, Prakash BA, Vreeken J (2016) Reconstructing an epidemic over time. In: Proceedings of the ACM conference on knowledge discovery and data mining. ACM, pp 1835–1844
49. Sapiezynski P, Stopczynski A, Lassen DD, Lehmann S (2019) Interaction data from the copenhagen networks study. *Sci Data* 6(1):1–10
50. Schwartz JT (1980) Fast probabilistic algorithms for verification of polynomial identities. *J ACM* 27(4):701–717
51. Schwikowski B, Uetz P, Fields S (2000) A network of protein-protein interactions in yeast. *Nat Biotechnol* 18(12):1257–1261
52. Sengupta N, Bagchi A, Ramanath M, Bedathur S (2019) ARROW: approximating reachability using random walks over web-scale graphs. In: Proceedings of the international conference on data engineering. IEEE, pp 470–481
53. Thejaswi S, Gionis A (2020) Pattern detection in large temporal graphs using algebraic fingerprints. In: Proceedings of the SIAM international conference on data mining. SIAM, pp 37–45
54. Thejaswi S, Gionis A, Lauri J (2020) Finding path motifs in large temporal graphs using algebraic fingerprints. *Big Data* 2020:335–362
55. Thejaswi S, Lauri J, Gionis A (2024) Restless-reachability-experimental-v1.0. <https://github.com/suhastheju/restless-reachability>
56. Thompson WH, Brantefors P, Fransson P (2017) From static to temporal network theory: applications to functional brain connectivity. *Netw Neurosci* 1(2):69–99
57. Tong H, Faloutsos C, Gallagher B, Eliassi-Rad T (2007) Fast best-effort pattern matching in large attributed graphs. In: Proceedings of the international conference on Knowledge discovery and data mining. ACM, pp 737–746
58. Von Christian F, Taras H, Holovatch Y, Palchykov V (2009) Public transport networks: empirical analysis and modeling. *Eur Phys J B* 68(2):261–275
59. Williams R (2009) Finding paths of length k in $O^*(2^k)$ time. *Inform Process Lett* 109(6):315–318
60. Wu H, Cheng J, Ke Y, Huang S, Huang Y, Wu H (2016) Efficient algorithms for temporal path computation. *Trans Knowl Data Eng* 28(11):2927–2942
61. Wu H, Huang Y, Cheng J, Li J, Ke Y (2016) Reachability and time-based path queries in temporal graphs. In: Proceedings of the international conference on data engineering. IEEE, pp 145–156
62. Zippel R (1979) Probabilistic algorithms for sparse polynomials. In: Proceedings of the international symposium on symbolic and algebraic computation (LNCS), Springer, vol 72, pp 216–226
63. Zschoche P (2023) Restless temporal path parameterized above lower bounds. In: Proceedings of the International symposium on theoretical aspects of computer science (LIPIcs), Schloss Dagstuhl-Leibniz-Zentrum für Informatik, Dagstuhl, Germany, vol 254, pp 55:1–55:16
64. Zschoche P, Fluschnik T, Molter H, Niedermeier R (2020) The complexity of finding small separators in temporal graphs. *J Comput Syst Sci* 107(2020):72–92

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.