Single source distance oracle for planar digraphs avoiding a failed node or link

Surender Baswana* Utkarsh Lath* Anuradha S. Mehta*

Abstract
Let $G = (V, E)$ be a directed planar graph on $n = |V|$ vertices, and let $s \in V$ be any fixed source vertex. We show that $G$ can be preprocessed in $O(n \text{polylog } n)$ time to build a data structure of $O(n \text{polylog } n)$ size which can answer the following query in $O(\log n)$ time for any $u, v \in V$ :

\[ \text{report distance from } s \text{ to } v \text{ in the graph } G \setminus \{u\} \]

We also address the all-pairs version of this problem and present a data structure with $O(n^{1/2} \text{polylog } n)$ preprocessing time and space which guarantees $O(n \text{polylog } n)$ query time.

1 Introduction
The problem of finding shortest paths in graphs is one of the most fundamental algorithmic problems of computer science. In the last few decades, researchers have explored the following variant of the shortest paths problem, called the replacement paths problem. Let $G = (V, E)$ be a directed weighted graph on $n = |V|$ vertices and $m = |E|$ edges. For any vertices $u, v, x \in V$, let $P(u, v, x)$ denote the shortest path from $u$ to $v$ in the graph $G \setminus \{x\}$. The aim is to have a compact data structure so that the path $P(u, v, x)$ and its length can be reported efficiently. In addition to being a problem of independent interest, the replacement paths problem appears to be a natural extension of the shortest paths problem. Efficient solution of this problem and its variants have been useful in dynamic routing as well.

The all-pairs version of the replacement paths problem (where $u, v$ and $x$ can be chosen arbitrarily) has been solved quite efficiently. Bernstein and Karger [3], improving the work of Demetrescu et al. [4], designed a data structure which can report length of $P(u, v, x)$ for any $u, v, x \in V$ in $O(1)$ time and, it takes $O(n^2 \text{polylog } n)$ space and $O(mn \text{polylog } n)$ preprocessing time. These bounds match, up to poly-logarithmic factors, the best bounds known for the classical all-pairs shortest paths problem. Duan and Pettie [6] addressed an even harder variant of the all-pairs replacement paths problem to handle failure of any two vertices.

For the single pair version of the replacement paths problem (where $u$ and $v$ are two fixed vertices and $x$ can be any vertex from $V$), no significant breakthrough has yet been made. For this problem, the trivial upper bound $O(mn)$ is still the best known upper bound, and there is a lower bound of $\Omega(mn)$ established by Hershberger et al. [11]. However, for various weaker versions of this problem, efficient results have been obtained in the past [1, 2, 15, 16, 17].

We address the replacement paths problem in planar graphs. Planar graphs have occupied the center stage of the classical shortest paths problem and many interesting and novel results have been achieved [5, 12, 13, 20]. Emek et al. [7] designed an $O(n \log^3 n)$ time solution for the single pair version of the replacement paths problem in planar graphs. Through a series of refinements [13, 21], this bound has been improved to $O(n \log n)$. The result has also been extended to bounded genus graphs by Erickson and Nayyeri [8].

We consider the single source version of the replacement paths problem: Given a directed weighted planar graph and a source vertex $s$, design a compact data structure to report $P(s, v, x)$ for any $v, x \in V$. As the central contribution of this paper, we present an extremely compact data structure for this problem. This data structure takes $O(n \text{polylog } n)$ space and can be built in $O(n \text{polylog } n)$ time. Interestingly, it takes just $O(\log n)$ time for reporting the length of $P(s, v, x)$ for any $v, x \in V$, hence the name “oracle”. Moreover, the entire path $P(s, v, x)$ can also be reported by spending $O(\log n)$ time per edge on the path. This data structure also leads to a near optimal algorithm for the single source second shortest paths problem in planar graphs.

We also obtain the first subquadratic space and sublinear query time solution for the all-pairs version of the replacement paths problem in planar graphs. Let $S \in [n^{3/2}, n^3]$. We present a data structure which occupies $O(S)$ space and can report $P(u, v, x)$ for any $u, v, x \in V$ in $O(n^2/S \text{polylog } n)$ time. This data structure is obtained, with effortless ease, by combining our data structure for the single source replacement paths problem with the data structure for all-pairs shortest paths devised by Djidjev [5]. We just provide its sketch in Appendix.

---

*Department of CSE, I.I.T. Kanpur, India. Email: {sabswana, utkarsh.lath, anuradha.s.mehta}@gmail.com. Research supported by the Indo-German Max Planck Center for Computer Science (IMPECS).
1.1 Overview of the result and new ideas

Let $T$ be the shortest paths tree rooted at source $s$ in $G$. Sleator and Tarjan [18] gave a technique for breaking down a tree into vertex disjoint paths. Using this technique in a straightforward manner, it suffices if we can solve the single source replacement paths problem where the failing vertex belongs to a given path $Q$ in tree $T$. So we focus on the latter problem and design a near-optimal data structure for it. At the heart of our data structure lies an efficient solution to a related problem, we call, \textit{Shortest Paths Avoiding Suffix Subpaths (SPASS)}. To build our data structure, we recursively partition $Q$ into equal halves, and build a balanced binary tree $H(Q)$ with $O(n)$ nodes. Each node $\nu$ of this tree corresponds to a subpath of $Q$. We augment $\nu$ with a data structure to solve the SPASS problem for its subpath. Interestingly, the replacement path $P(s, v, x)$, for any $v \in V, x \in Q$, can be computed by querying SPASS data structure associated with a single node in $H(Q)$.

It turns out that SPASS can be reduced to the Multiple Source Shortest Paths (MSSP) problem. Klein [12] designed an $O(n \log n)$ space and time algorithm to solve MSSP in planar graphs. This by itself would complete the solution to our problem, but solving $O(n)$ instances of the SPASS problem on the given graph would take $O(n^2)$ space and time. To circumvent this problem, we build a relatively small sized graph for each node of $H(Q)$ and show that it suffices to make SPASS queries on these graphs only. To build these small size graphs, we proceed as follows.

1. We make a couple of insightful observations about SPASS data structures associated with various nodes of $H(Q)$. In particular, SPASS data structure at a node bears some similarity with SPASS data structure of some of its ancestor. This similarity allows us to obtain a small sized graph for a node by \textit{compressing} the subgraph associated with some of its ancestor. This compression involves a careful pruning and shortcutting paths by edges.

2. We discover an interesting property about replacement paths in planar graph (Theorem 3.1). This property, which might be of independent interest, proves to be helpful for our problem in the following manner. In order to reduce the size of graph associated with a node in $H(Q)$ without destroying planarity, we need to add some edges which are \textit{short-cuts} of certain paths. Note that, addition of edges, in general, may lead to creation of erroneous replacement paths in a graph. The property (Theorem 3.1) ensures that no such erroneous path gets created if these edges are assigned weights \textit{suitably}.

Organization of the paper

After introducing various notations in section 2, we describe an important property of replacement paths in planar graphs in section 3. Section 4 and 5 present efficient solutions of a restricted version of the single source replacement paths problem wherein the failing vertex belongs to a given path. This forms the core of our main result. We use this in conjunction with a divide and conquer strategy to solve the single source replacement paths problem in section 6.

2 Preliminaries

Let $G = (V, E)$ be the given planar directed graph and $\omega : E \rightarrow \mathbb{R}^+$ be a weight function on its edges. Let $s \in V$ be a source vertex and $T$ be the shortest path tree in $G$ rooted at vertex $s$. Without loss of generality, we assume uniqueness of shortest paths and replacement paths. Throughout this paper, we will be working with that planar embedding of $G$ which has $s$ lying on the boundary of the infinite face. Let $Q$ be any path in $G$. We define the following notations.

- $T(u)$: the subtree of $T$ rooted at $u$.
- $Q(u, v)$: the subpath of $Q$ from vertex $u$ to vertex $v$. It is defined only if $u, v \in Q$, and $u$ precedes $v$.
- $P(u, v)$: the shortest path from $u$ to $v$ in $G$.
- $\delta(u, v)$: the length of $P(u, v)$.
- $P(u, v, Q)$: the shortest path from $u$ to $v$ in $G$ when all the vertices of path $Q$ fail.
- $\delta(u, v, Q)$: the length of $P(u, v, Q)$.
- $\text{FIRST}(Q)$: the first vertex of $Q$.
- $\text{LAST}(Q)$: the last vertex of $Q$.

\textbf{Definition 2.1. (family)}: Given a path $Q$ in $T$, let $R$ be a subpath of $Q$. Let $x$ be the vertex immediately after $\text{LAST}(R)$ on $Q$. We define $\text{FAMILY}(R, Q)$ (see Figure 1) as the set of vertices present in $T(\text{FIRST}(R)) \setminus T(x)$.

![Figure 1: The vertices in the shaded region constitute FAMILY(R, Q) where R = Q(u, v).](image)

\textbf{Definition 2.2. (G(x))}: The subgraph of $G$ induced by vertices of $T(x)$ and augmented by vertex $s$ and edges from $s$ as follows. For each $v \in T(x)$ with neighbors outside $T(x)$, add an edge $(s, v)$ with weight $= \min_{u \in E, u \in T(x)} (\delta(s, u) + \omega(u, v))$.

It can be observed that we have to just work on $G(x)$ instead of $G$ if we want to handle failure of any vertex.
belonging to \( T(x) \). Notice that a replacement path possesses optimal substructure property. This leads to the following useful observation.

**Observation 2.1.** Let \( v \) be a vertex and let vertex \( x \in P(s, v) \). If \( P(s, v, x) \) passes through a vertex \( w \in P(x, v) \), then \( P(s, v, x) = P(s, w, x) + P(w, v) \).

### 3 Bounded difference property of replacement paths

Let \( v, w \in V \) be any two vertices and \( x \) be a failing vertex. It is easy to observe that the difference between the lengths of the two replacement paths \( P(s, v, x) \) and \( P(s, w, x) \) in a general graph may be quite arbitrary. However, for planar graphs, this difference is bounded if \( v \) and \( w \) satisfy certain topological conditions.

Let \((u, v)\) be an edge of \( G \) that is not present in tree \( T \). Let \( a \) be the lowest common ancestor of \( u \) and \( v \) in tree \( T \). Notice that \( P(a, u) \), edge \((u, v)\), and \( P(a, v)\) together form a separator in \( G \). See Figure 2(i). In fact, ignoring the directedness of the graph, this separator is actually the fundamental cycle defined by \((u, v)\) and tree \( T \). Let \( w \) be any vertex such that \( s \) and \( w \) belong to different regions of the plane defined by this cyclic separator. The bounded-difference-property of replacement paths in planar graphs claims the following: No matter which vertex fails, the replacement path to \( w \) can not be arbitrarily longer than the replacement path to \( v \). This claim holds even when an entire subpath of \( T \) (not just a vertex) fails. The following theorem states this property in a formal manner.

**Theorem 3.1.** Consider an embedding of the fundamental cycle defined by any non-tree edge \((u, v)\) and tree \( T \). Let \( w \) be any vertex such that \( s \) and \( w \) lie on different sides of this cycle. Then for any path \( Z \) in the tree \( T \) with \( s, v, w \notin Z \),

\[
(3.1) \delta(s, v, Z) - \delta(s, w, Z) \leq \delta(s, u) + \omega(u, v) - \delta(s, w)
\]

**Proof.** Let \( a \) be the lowest common ancestor of \( u \) and \( v \) in \( T \) (see Figure 2(ii)). If \( P(s, a) \) remains intact after failure of \( Z \) then Equation 3.1 holds obviously. So let us assume that the failing path \( Z \) includes a subpath of \( P(s, a) \) (see Figure 2(ii)). Notice that the fundamental cycle defined by \((u, v)\) is a separator with \( s \) and \( w \) lie on two different sides of it. So the path \( P(s, w, Z) \) must intersect this separator at some vertex, say \( y \). There are only two possibilities of this path as shown in Figure 2(ii). Firstly observe that

\[
(3.2) \delta(s, y) + \delta(y, v, Z) \leq \delta(s, u) + \omega(u, v)
\]

Now using triangle inequality, we can infer that
\[
\delta(s, w) \leq \delta(s, y) + \delta(y, w, Z) \quad \text{and} \quad \delta(s, v, Z) \leq \delta(s, y, Z) + \delta(y, v, Z).
\]

Combining together these inequalities and using \( \delta(s, w, Z) = \delta(s, y, Z) + \delta(y, w, Z) \) we get

\[
\delta(s, w) + \delta(s, v, Z) \leq \delta(s, w, Z) + \delta(y, v, Z)
\]

Plugging the inequality from Equation 3.2, RHS is bounded by \( \delta(s, w, Z) + \delta(s, u) + \omega(u, v) \). Rearranging the terms, we get the desired Equation 3.1.

### 4 Data structure for failure of vertices on a path \( Q \subseteq T \)

Let \( Q \) be any path in \( T \) originating from source \( s \). In this section we describe a data structure to report the replacement path from source \( s \) to any vertex in \( G \) when the failing vertex lies on \( Q \). This data structure will be based on a problem called Shortest Paths Avoiding Suffix Subpaths, or SPASS in short.

#### 4.1 spass problem

A subpath \( R \) of a path \( Q \) is called a suffix subpath of \( Q \) if \( \text{LAST}(R) = \text{LAST}(Q) \).

**Definition 4.1.** (The spass problem) Given a shortest path \( Q \) in \( T \), design a compact data structure for efficiently reporting shortest path from \( s \) to any vertex \( v \in V \) avoiding (i.e. not passing through any vertex of) any given suffix subpath of \( Q \).

There is no known efficient solution of SPASS for general graphs. However, for planar graphs, there exists a solution of SPASS with \( O(|G| \text{FIRST}(Q) + \log n) \) space and \( O(\log n) \) query time. This solution is obtained by a simple reduction to the Multiple Source Shortest Paths (MSSP) problem, for which Klein [12] gave an efficient algorithm. We describe the reduction of SPASS to MSSP in Section 5. Note that Emek et al. [7] and Wulff-Nilsen [21] also used reduction to MSSP for solving similar subproblems in their algorithms for single pair replacement paths problem.

#### 4.2 Handling failure of a vertex from \( Q \)

We begin with showing a nice relationship between our replacement paths problem and SPASS problem. Let \( Q(u, v) \) be a subpath of \( Q \), and consider the failure of any vertex \( x \in Q(u, v) \). Let \( w \) be any descendant of \( v \) in \( T \). The replacement path \( P(s, w, x) \) would either skip...
the entire suffix of $Q(u, v)$ starting from $x$ or it definitely passes through $v$ (see Figure 3).

\[ \text{Figure 3: Two cases of } P(s, w, x) \text{ (shown thick) from perspective of subpath } Q(u, v). \]

In the first case, the replacement path to $w$ can be computed by a single spass query on $Q(u, v)$. In the second case, using Observation 2.1, $P(s, w, x)$ can be seen as the concatenation of the replacement path $P(s, v, x)$ and the shortest path $P(v, w)$ already available in $T$. Based on these ideas, we now present a data structure $\mathcal{H}(Q)$ to solve replacement paths problem when the failure vertex lies on $Q$.

$\mathcal{H}(Q)$ is a balanced binary tree. The root node of $\mathcal{H}(Q)$ is associated with path $Q$ itself. If $Q$ has only 1 vertex, then $\mathcal{H}(Q)$ is a singleton tree. If $Q$ has more than 1 vertex, we divide $Q$ into 2 subpaths $Q_1$ and $Q_2$ with (nearly) equal number of vertices. So, $Q = Q_1 + Q_2$. In this case, $\mathcal{H}(Q)$ is a binary tree with $\mathcal{H}(Q_1)$ and $\mathcal{H}(Q_2)$ as its left child and right child respectively.

Let $\nu$ be any node in tree $\mathcal{H}(Q)$. Throughout the paper, we shall use $\nu$ synonymously with the subpath of $Q$ associated with $\nu$. We will denote the left child of $\nu$ in $\mathcal{H}$ by $\nu.L$, and its right child by $\nu.R$. See Figure 4 for a better understanding of $\mathcal{H}(Q)$ and these concepts.

\[ \text{Figure 4: Binary tree } \mathcal{H}(Q) \text{ constructed on path } Q. \]

Each node $\nu$ of $\mathcal{H}(Q)$ is augmented with the following information.

\begin{itemize}
  \item [(a)] A data structure $\nu.spass$ which solves spass problem on $\nu.L$. In particular, a query $\nu.spass(w, x)$, for any $x \in \nu.L$ and any $w \in \text{family}(\nu, Q)$, will report the length of the shortest path from $s$ to $w$ avoiding the entire suffix of $\nu.L$ starting from $x$.\[ \text{distance from } s \text{ to the last vertex of } \nu.L \text{ avoiding } x. \]
  \item [(b)] For each $x \in \nu.L$, store $\delta(s, \text{Last}(\nu.L), x)$, i.e., the range $\delta(s, \text{Last}(\nu.L), x) = \delta(s, \text{Last}(\nu.L), x) + \delta(\text{Last}(\nu.L), w)$
\end{itemize}

We now describe an algorithm to report $\delta(s, w, x)$ for any failing vertex $x \in Q$ and destination vertex $w$. Without loss of generality assume that $w$ is descendant of $x$ in $T$ (otherwise distance to $w$ remains intact even after the failure of $x$). We introduce a notation here. Let $\text{path-parent}(w, Q)$ denote the vertex of $Q$ at which $P(s, w)$ leaves $Q$; it is the vertex of in Figure 3. We first compute the Lowest Common Ancestor (LCA) of $\text{path-parent}(w, Q)$ and $x$ in $\mathcal{H}(Q)$. Let it be $\nu$. The distance $\delta(s, w, x)$ can be retrieved from the data structure associated with $\nu$ as described by Algorithm 1. This algorithm is based on the ideas discussed above. In particular, the else part is captured precisely in Figure 3 with $\nu.L$ being the subpath $Q(u, v)$.

Let us analyse the running time of Algorithm 1. Notice that $\text{path-parent}(w, Q)$ is the lowest common ancestor of $w$ and $\text{Last}(Q)$ in tree $T$. Algorithm 1 makes two LCA queries - one on $T$ and one on $\mathcal{H}(Q)$, which take $O(1)$ time (see [10]). In addition, it makes a single spass query which takes $O(\log n)$ time. Hence Algorithm 1 takes $O(\log n)$ time.

\begin{algorithm}
\caption{Reporting $\delta(s, w, x)$ where $x$ is a failing vertex on path $Q$}
\begin{algorithmic}
\STATE $\nu \gets \text{lca}(x, \text{path-parent}(w, Q));$
\IF $(x = \text{Last}(\nu.L))$
\STATE return $\nu.spass(w, x)$
\ELSE
\STATE return $\{ \nu.spass(w, x), \delta(s, \text{Last}(\nu.L), x) + \delta(\text{Last}(\nu.L), w) \}$
\ENDIF
\end{algorithmic}
\end{algorithm}

Let us discuss preprocessing of the data structure $\mathcal{H}(Q)$. Assuming that for each node of $\mathcal{H}(Q)$, part (a) has already been computed, we can compute part (b) for all nodes efficiently by pursuing a bottom-up approach as follows. Consider any node $\mu \in \mathcal{H}(Q)$. As part (b) for $\mu$, we need to compute $\delta(s, \text{Last}(\mu.L), x)$ for all $x \in \mu.L$. If we invoke Algorithm 1 with $w = \text{Last}(\mu.L)$, it will query node $\nu = \text{lca}(x, \text{Last}(\mu.L))$. Notice that $\nu$ is descendant of $\mu$ in $\mathcal{H}(Q)$ and we have been pursuing a bottom-up approach. Therefore, the node $\nu$ is already equipped with part (a) and (b) of its information. Hence it will take $O(\log n)$ time to compute $\delta(s, \text{Last}(\mu.L), x)$ for any $x \in \mu.L$. Thus all the nodes of $\mathcal{H}(Q)$ can be augmented with part (b) in $O(n \log^2 n)$ time. Leaving this overhead of $O(n \log^2 n)$, the main task in preprocessing of $\mathcal{H}(Q)$ is the computation of part (a), that is, $\nu.spass$ for each $\nu \in \mathcal{H}(Q)$. If we naively use the MSSP reduction to compute and store $\nu.spass$ for each $\nu \in \mathcal{H}$, then
ν.spass will require $O(|G(\text{FIRST}(\nu))| \log n)$ space and preprocessing time. This will lead to $\Theta(n^2)$ space and preprocessing time for the entire $\mathcal{H}(Q)$. In the following section, we use novel insights to show that spass data structures at deeper nodes in $\mathcal{H}(Q)$ can be built on smaller proxy graphs instead of the original input graph $G$. This leads to $O(n \log^3 n)$ bound on the space and the preprocessing time of $\mathcal{H}(Q)$.

5 Efficiently building the data structure $\mathcal{H}(Q)$
We begin with an overview of the mssp algorithm of Klein [12] followed by the reduction of spass to mssp.

5.1 Overview of mssp algorithm of Klein [12]
Given a directed planar graph $G$, and a face $f$ in $G$, mssp problem aims at preprocessing a graph that can efficiently answer queries of the type: report $\delta(s, t)$ for any $s \in f, t \in V$. Klein [12] designed an $O(n \log n)$ time algorithm to build an $O(n \log n)$ size data structure for the mssp problem. The query time achieved is $O(\log n)$. Before giving an overview of this algorithm, we begin with some well known terminologies related to shortest path problem.

Let $T$ be a tree rooted at a vertex $s$ in $G$, and let $d_T(v)$ denote the distance from $s$ to vertex $v$ in this tree. An edge $e = (x, y) \in E \setminus T$ is called tense if $d_T(y) > d_T(x) + \omega(e)$. Tree $T$ will be the shortest paths tree iff no (non-tree) edge is tense. Relaxation of a tense edge $e$ in a tree $T$ involves removing the edge incident to $y$ in $T$ and adding $e$ to $T$. Hence, the task of transforming any tree $T$ into the shortest path tree can be achieved through a sequence of relaxations of the tense edges.

We now present a high-level description of mssp algorithm of Klein [12]. Enumerate all vertices on $f$ as $s_1, s_2, \ldots, s_{|f|}$ in anti-clockwise direction. The algorithm starts with the shortest path tree $T_1$ rooted at $s_1$ and it constructs the rest of the trees $T_i$, $i > 1$ incrementally and implicitly. Given a shortest path tree $T_{i-1}$, the tree $T_i$ is built in the following manner: Consider the source vertex making transition from $s_{i-1}$ to $s_i$ in $T_{i-1}$. As a result, there will be some non-tree edges in the graph which become tense. Carrying out relaxation until there is no tense edge left will provide $T_i$. The mssp algorithm of Klein [12] computes the tense edges and carries out their relaxation in a specific order such that exactly those edges get relaxed which belong to $T_i \setminus T_{i-1}$. In this manner $T_{i-1}$ gets transformed to $T_i$ (see Figure 5).

Klein [12] exploited planarity crucially to show that during the entire movement of source from $s_1$ to $s_{|f|}$, every edge in the graph is relaxed at most once. All the trees $T_i$’s can thus be expressed by the shortest path tree $T_1$ and a sequence of $O(n)$ edge relaxations as the source moves from $s_1$ to $s_{|f|}$. If an edge $e$ gets relaxed while computing shortest path $T_i$ from $T_{i-1}$, we record the relaxation of $e$ as a pair $(e, i)$ and call $i$ as relaxation index of $e$. The sequence of edge relaxations is processed to build an $O(n \log n)$ space persistent data structure which implicitly stores all trees $T_i$. This data structure takes $O(\log n)$ time to report the distance of any vertex $v \in G$ from any vertex $s_i \in f$.

5.1.1 Reduction of spass to mssp
We now give a reduction from spass to mssp. Let $Q = (s = x_0, x_1, \ldots, x_\tau = t)$ be a given shortest path in $T$. Given an edge $e = (x_i, y)$ with $e \notin Q$, we say that $e$ lies to the left (similarly right) of $Q$ if $i \in \{0, \tau\}$ or $e$ occurs strictly between $(x_i, x_{i+1})$ and $(x_{i-1}, x_i)$ in the counterclockwise (clockwise) order. Let $R$ be the replacement path from $s$ to a vertex $v$ when some suffix subpath $Q(x_j, t)$ of $Q$ fails. We say that $R$ leaves $Q$ from left (similarly from right) if the first edge on $R$ which is not shared by $Q$ lies to the left (right) of $Q$. We build data structure spass, to retrieve the paths $R$ that leave $Q$ from the left, by a reduction to mssp problem as given below. A similar data structure spass, can be built to retrieve paths $R$ that leave $Q$ from the right.

We build a graph $G'$ by modifying $G$ as follows (see Figure 6):

- Reverse the direction of all edges on the path $Q$ and assign zero weight to each of them.
- For each vertex $x_i \in Q$, increase the weight of all its outgoing edges which lie to the left of $Q$ by $\delta(s, x_i)$.
- Remove all other edges incident to vertices of $Q$.
- Add edge $(s, t)$ of weight $\infty$. This creates a face $f$
which has all the vertices of \( Q \) on it.

Lemma 5.1 states the relationship between SPASS on path \( Q \) in \( G \) and MSSP on face \( f \) in \( G' \).

**Lemma 5.1.** Length of the shortest path from \( s \) to \( v \) avoiding \((x_{i+1}, \ldots, x_r)\) in \( G \) which leaves \( Q \) from left is the same as the length of the shortest path from \( x_i \) to \( v \) in \( G' \).

Because of this direct relation between SPASS and MSSP we use them interchangeably in the later sections.

### 5.2 Ideas leading to linear space and preprocessing time for \( \mathcal{H}(Q) \)

For the sake of compactness, in the following discussion, we use \( \mathcal{H} \) to denote \( \mathcal{H}(Q) \), \( \text{FAMILY}(\nu) \) to denote \( \text{FAMILY}(\nu, Q) \), and \( G(\nu) \) to denote \( G(\text{FIRST}(\nu)) \).

We introduce a notation here. For a node \( \nu \in \mathcal{H} \), define L-PARENT(\( \nu \)) as the nearest ancestor \( \mu \) of \( \nu \) in \( \mathcal{H} \) such that \( \nu \) appears in the left subtree of \( \mu \). Conversely, \( \nu \) is said to be a L-CHILD of \( \mu \). See Figure 7.

![Figure 7: A node \( \mu \) and its L-CHILDREN in \( \mathcal{H} \).](image)

Notice that \( \nu \) is a suffix subpath of \( \mu.L \) and \( \text{FAMILY}(\nu) \) is a subset of \( \text{FAMILY}(\mu) \). As a result, \( \text{SPASS}(\mu) \) and \( \text{SPASS}(\nu) \) have certain common features as follows. Consider any vertex \( w \in \text{FAMILY}(\nu) \) and \( x \in \nu.L \) and let us analyse the structure of the path \( \nu.\text{SPASS}(w, x) \). Note that \( \mu.\text{SPASS}(w, x) \) avoids entire suffix of \( \mu.L \) starting from \( x \). However, \( \nu.\text{SPASS}(w, x) \) can potentially use some portion of \( \mu.L \); in particular, it can use \( \nu.R \). This slight distinction between \( \nu \) and \( \mu \) leads us to the following observations. If the path \( \nu.\text{SPASS}(w, x) \) does not use any vertex of \( \nu.R \), then it is the same as \( \mu.\text{SPASS}(w, x) \) (see Figure 8 (i)). If not, let \( \nu \) be the first vertex of the path \( \nu.\text{SPASS}(w, x) \) which lies on \( \nu.R \). One possibility is that the subpath from \( v \) to \( w \) uses only vertices of \( \text{FAMILY}(\nu) \). This is shown in Figure 8 (ii). Otherwise, let \( a \) be the last vertex of the path that does not belong to \( \text{FAMILY}(\nu) \). At first glance, it may appear that the portion of \( \nu.\text{SPASS}(w, x) \) between \( v \) and \( a \) may be quite arbitrary. However, notice that the path \( P(v, a) \), which is already in \( T \), is intact even when any suffix subpath of \( \nu.L \) fails. Therefore, using Observation 2.1 the subpath of \( \nu.\text{SPASS}(w, x) \) from \( v \) to \( a \) is just \( P(v, a) \). Figure 8 (iii) depicts this case.

Considering all these cases, observe that any \( \nu.\text{SPASS} \) query can be answered by building SPASS data structure on a graph which, in addition to having all vertices and edges of \( \text{FAMILY}(\nu) \), satisfies the following properties.

1. For all vertices \( w \in \text{FAMILY}(\nu) \setminus \nu.R \), it has the path \( \mu.\text{SPASS}(w, x) \) for all \( x \in \nu.L \). This will take care of \( \nu.\text{SPASS}(w, x) \) shown in Figure 8 (i).
2. For vertices \( w \in \nu.R \), it has the path \( \mu.\text{SPASS}(w, x) \) for all \( y \) such that \( (y, w) \) is an edge in \( G \). Note that \( y \) may not necessarily be present in \( \text{FAMILY}(\nu) \). This property, along with property 1, will take care of \( \nu.\text{SPASS}(w, x) \) shown in Figure 8 (ii).
3. For all edges \( (a, b) \) such that \( a \notin \text{FAMILY}(\nu) \) and \( b \in \text{FAMILY}(\nu) \), the graph has a path of length \( \delta(\text{LAST}(\nu), a) + \omega(a, b) \) from \( \text{LAST}(\nu) \) to \( b \). This property, along with property 1 and 2, will take care of \( \nu.\text{SPASS}(w, x) \) shown in Figure 8 (iii).

For each \( \nu \), we construct a small sized graph \( G^{\text{reduced}}(\nu) \), or \( G'(\nu) \) in short, which satisfies the above three properties. Before providing the complete details of construction of \( G'(\nu) \), we provide a sketch here. We start with \( G'(\mu) \) and sparsify it by keeping only those edges which are either incident to some vertex of the set \( \text{FAMILY}(\nu) \) or belong to some path \( \mu.\text{SPASS}(w, x) \), \( x \in \nu.L \). This graph will satisfy the first two properties mentioned above. We then reduce its size by removing some vertices and shortcutting some paths by edges. We call this graph \( G'(\nu) \). Thereafter, we modify \( G'(\nu) \) and construct \( G'(\nu) \) so that it also satisfies the third property. A trivial way to construct \( G'(\nu) \) from \( G'(\nu) \) would be the following. For every edge \( (a, b) \) such that \( a \notin \text{FAMILY}(\nu) \) and \( b \in \text{FAMILY}(\nu) \), add an edge from \( \text{LAST}(\nu) \) to \( b \) of required length. However, if we do so, the resulting graph may no longer be planar. To overcome this problem, we employ the bounded difference property of replacement paths in planar graphs (Theorem 3.1).

#### 5.3 Construction of \( G'(\nu) \)

We construct the graphs \( G'(\nu) \) in a top-down manner on the tree \( \mathcal{H} \). The nodes of \( \mathcal{H} \) which have no L-PARENT are the ones that represent some suffix subpath of \( Q \). If \( \nu \) is one such node, then we set \( G'(\nu) = G(\nu) \). For example, for \( \nu = Q \), \( G'(Q) = G \). For other nodes \( \nu \) we assume that the three properties mentioned in the previous section hold for \( G'(\mu) \), where \( \mu = \text{L-PARENT}(\nu) \). We then construct \( G'(\nu) \) using \( G'(\mu) \), ensuring that the properties hold for \( G'(\nu) \). Thus, by induction, we ensure that for all \( \nu \in \mathcal{H} \), \( G'(\nu) \) satisfies the required properties.
5.3.1 Construction of $G^*(\nu)$ We require $G^*(\nu)$ to satisfy the first two properties mentioned above. Though $G^*(\mu)$ can serve as $G^*(\nu)$, we need a smaller subgraph for our purpose. Let $T^\mu(x)$ denote the shortest path tree rooted at $s$ in the graph $G^*(\mu) \setminus R$, where $R$ is the suffix subpath of $\mu.L$ starting at $x$. It is sufficient to have the edges of the set $\cup_{x \in \nu.L} T^\mu(x)$ only. These edges can also be viewed as: the edges of $T^\mu(\text{FIRST}(\nu))$ and the edges that get relaxed as the source moves from FIRST(\nu) to LAST(\nu.L) while executing MSSP on $G^*(\mu)$.

Let us assign red or blue color to the edges of $\cup_{x \in \nu.L} T^\mu(x)$. Assign red color to all those edges which have at least one end-point in FAMILY(\nu). Also assign red color to all those edges that are relaxed when the source moves from FIRST(\nu) to LAST(\nu.L) during MSSP on $G^*(\mu)$. The remaining edges are colored blue. We assign colors to the vertices as follows. The end-points of all red edges are colored red and the remaining vertices are colored blue.

5.3.2 Transforming $G^*(\nu)$ into $G^*(\nu)$ We first introduce some notations. An edge $e = (a, b)$ is a crossing edge if $a \notin \text{FAMILY}(\nu)$ and $b \in \text{FAMILY}(\nu)$, and a reverse crossing edge if $a \in \text{FAMILY}(\nu)$ and $b \notin \text{FAMILY}(\nu)$. Consider an imaginary line $L$ passing through LAST(\nu) which splits every crossing and reverse crossing edge. $L$ will act as a separator for $G^*(\nu)$, separating the vertices of FAMILY(\nu) from those that lie beyond family (see Figure 10). Starting from vertex $v_0 = \text{LAST}(\nu)$, we traverse along $L$ in any one direction (the same will be done in the other direction as well). Suppose we encounter edges $e_1 = (a_1, b_1), e_2 = (a_2, b_2), \ldots$, in that order, during the traversal. For each edge $e_i$, do the following:

- At the point of intersection of $e_i$ with $L$, create a new vertex $v_i$ and add an edge from $v_{i-1}$ to $v_i$. Remove the edge $(a_i, b_i)$ and add edges $(a_i, v_i)$ and $(v_i, b_i)$.
- Assign weights to the new edges as follows.
  (i) $\omega(v_{i-1}, v_i) = \delta(s, a_i) - \delta(s, a_{i-1})$.
  (ii) $\omega(a_i, v_i) = 0$.
  (iii) $\omega(v_i, b_i) = \omega(e_i)$.

It can be seen that the vertices that belong to or are neighboring to FAMILY(\nu) will be colored red. So in order to compute $G^*(\nu)$ of small size, we modify $\cup_{x \in \nu.L} T^\mu(x)$ such that distance between any two red vertices is preserved, but all blue vertices are removed. Notice that each blue edge is an edge of the tree $T^\mu(\text{FIRST}(\nu))$. So we first remove all blue edges and then connect every red vertex with its nearest red ancestor in the tree $T^\mu(\text{FIRST}(\nu))$ (if not already connected) by a new blue edge whose length is equal to the distance between them in the tree. After this procedure, the blue vertices are isolated and can be removed. See Figure 9.

![Figure 8: Possible structures of the path $\nu.SPASS(w, x)$.

![Figure 9: Transformation of $T^\mu(\text{FIRST}(\nu))$ to remove all blue vertices (shown hollow). As shown in (ii), only red vertices (shown solid) remain. The thick edges are red edges, the thin edges are blue edges.

![Figure 10: Transforming $G^*(\nu)$ into $G^*(\nu)$ by adding new vertices along (reverse) crossing edges. The dashed edge shows the false path added between $a_1$ and $b_2$.](image)
new path from last(ν) to b₁ will be δ(last(ν), a₁) + ω(e₁). Hence G*(ν) satisfies the third property mentioned earlier. Furthermore, if e₁ is a reverse crossing edge, the new path from last(ν) to b₁ will have length no less than δ(s, b₁) − δ(s, last(ν)) = δ(last(ν), b₁).

However, a problem with this transformation is that it introduces some “false paths” in the resulting graph: For every pair of edges e₁, e₂, l₁ < l₂, now there is a new path from a₁ to b₁ which was not present in the original graph (see Figure 10). We have to ensure that this “false path” does not affect the shortest path avoiding suffix path from s to the vertex b₁. The bounded difference property of replacement paths (Theorem 3.1) proves to be crucial here as follows.

Notice that in G*(ν), e₁ is not present in the shortest path tree rooted at s, and P(s, a₁), e₂, and P(s, b₁) form a separator in G*(ν) that separates s and a₁. Thus, by bounded difference property of replacement paths (Theorem 3.1), for any path Z in T not containing s and b₁,

δ(s, b₁, Z) − δ(s, a₁, Z) ≤ δ(s, a₁) + ω(e₁) − δ(s, a₁)

It can be verified that the RHS of the above inequality is the same as the length of the “false path” introduced after a₁ and b₁. Hence, the presence of the “false path” in G*(ν) will not alter δ(s, b₁, Z) for any Z.

5.4 Analysis We first provide a bound on \( \sum_{ν ∈ H} \sum_{v ∈ \text{FAMILY}(ν)}\deg(v) \). Any vertex v will belong to family of at most one node at any level of H because whenever a node ν ∈ H is partitioned into two halves L and R, every vertex v ∈ \text{FAMILY}(ν) goes either into the family of ν.L or into family of ν.R, but not both. There are O(log n) levels in H, hence v belongs to family of at most log n nodes. Therefore,

\[
\sum_{ν ∈ H} \sum_{v ∈ \text{FAMILY}(ν)}\deg(v) ≤ log n \sum_{v ∈ G}\deg(v) = O(n \log n)
\]

Now, we provide a bound on the sum of sizes of G*(ν) for all ν ∈ H. Consider any edge of G*(ν). It is of one of the 4 types:

1. A red edge that was incident to some vertex v ∈ \text{FAMILY}(ν). The number of such edges is at most the sum of degrees of all the vertices of \text{FAMILY}(ν). Such edges will be called red edges of type 1. The total number of such edges summed over all G*(ν) is O(n log n) due to Equation 5.3.

2. A red edge that was not incident to any vertex of \text{FAMILY}(ν). This edge must be one of those edges that were relaxed when the source moved from \text{FIRST}(ν) to \text{LAST}(ν.L) during MSSP on G*(µ) where µ = L-PARENT(ν). We call it red edge of type 2.

3. A blue edge. Note that at most one blue edge will be incident to any vertex. Moreover, any blue edge is between two red vertices, therefore, we can safely say that number of blue edges is no more than number of red edges. We will ignore counting blue edges henceforth.

4. An edge that was added when G*(ν) was constructed from G*(µ). Only 3 new edges are added for every edge whose one endpoint lies in \text{FAMILY}(ν). The number of such edges can be bounded by O(n \log n) due to Equation 5.3. These edges will also be called as red edges of type 1.

We only need to bound the total number of red edges of type 2. Any red edge of type 2 in G*(ν) must be present in G*(µ). If it was a red edge of type 2 in G*(µ), then it must have been passed from µ’s L-PARENT. Otherwise, it was either a blue edge or a red edge of type 1 in G*(µ). We can ensure that no blue edge of G*(µ) gets passed down from µ to ν as a red edge of type 2 in G*(ν) (see Appendix for details). In that case, every red edge of type 2 in G*(ν) must have been a red edge of type 1 in a graph G*(λ) for some ancestor λ of ν; and it must have been “passed on” from there to G*(ν) through a series of “L-ANCESTORS” of ν such that it was a red edge of type 2 at all those ancestors. This suggests that for each red edge of type 1 at a node in H, we need to count the number of its descendants in H where it appears as a red edge of type 2.

Consider a node µ ∈ H and let ν₁, ν₂, . . . , νₖ be its L-CHILDREN (see Figure 7). Consider any red edge e in G*(µ). Recall that while executing MSSP on G*(µ), an edge gets relaxed at most once. Therefore, if x ∈ µ.L be such that e got relaxed while computing T*(x) and x ∈ νₙ.L, then e will be present as a red edge of type 2 only in G*(νₙ). Hence, any red edge of type 1 present in G*(µ) will be passed on to at most one of its L-CHILDREN as a red edge of type 2. Also, this L-CHILD will in turn pass it on to at most one of its L-CHILDREN and this sequence will go on till we reach a leaf of the tree H. So any red edge of type 1 present at µ will be present as a red edge of type 2 at no more than log n nodes of H. Hence, the total number of red edges of type 2, summed over all nodes of H will be at most log n times the total number of red edges of type 1 summed over all nodes of H. Hence, \( \sum_{ν ∈ H} |G*(ν)| = O(n \log² n) \). The spss data structure on a n vertex graph occupies \( O(n \log n) \) space, so the total space requirement of data structure associated with H(Q) is \( \sum_{ν ∈ H} O(|G*(ν)| \log n) = O(n \log³ n) \). Building spss on n vertex graph takes \( O(n \log n) \) time. G*(ν) can be built from G*(µ) in \( O(|G*(µ)|) \) time. However, a node µ is L-PARENT of at most log n nodes of H, so the
preprocessing time of $\mathcal{H}(Q)$ is:

$$\sum_{\nu \in \mathcal{H}, \mu = \text{L-PARENT}(\nu)} O(|G'(\nu)| \log n + |G'(\mu)|) = O(n \log^3 n)$$

We can thus state the following theorem.

**Theorem 5.1.** For any path $Q$ present in the shortest path tree $T$ rooted at vertex $s$ in $G$, a data structure of $O(n \log^3 n)$ size can be built in $O(n \log^3 n)$ time that can report $\delta(s, v, x)$ in $O(\log n)$ time for any $x \in Q, v \in V$.

### 6 Data structure for the single source replacement paths problem

In order to solve the single source replacement paths problem, we use Theorem 5.1 and pursue a divide and conquer approach. We compute a path $Q$ in tree $T$ using heavy path decomposition technique given by Sleator and Tarjan [18] as follows. Starting from $s$, we traverse down the tree $T$ and keep extending $Q$ along that edge from which the largest (in terms of no. of nodes) subtree hangs. We stop when we reach a leaf node. The path $Q$ created in this manner has the following nice property.

Let $v_1, \ldots, v_k$ be the roots of the subtrees of $T$ connected to the path $Q$ with an edge. Each of $T(v_i)$’s are disjoint and each $T(v_i)$ has fewer than $n/2$ vertices.

First we build a data structure for handling failure of any vertex from $Q$ according to Theorem 5.1. Then, for each $1 \leq i \leq k$, we recursively solve the single source replacement paths problem for $G(v_i)$ (see Definition 2.2). It follows from the property mentioned above that $G(v_i)$ has at most $n/2$ vertices. So the depth of the recursion will be $O(\log n)$ only. We can thus state the following theorem.

**Theorem 6.1.** Given a directed planar graph $G = (V, E)$ on $n$ vertices and a source vertex $s$, a data structure of $O(n \log^4 n)$ size can be built in $O(n \log^4 n)$ time which can report $\delta(s, v, x)$ in $O(\log n)$ time for any $v, x \in V$.

### 7 Conclusion

We presented a single source distance oracle for planar digraphs avoiding any single failed vertex. Its size and construction time are optimal up to poly-logarithmic factors. It can be adapted to handle edge failure as well by introducing a vertex at the center of each edge. We would like to conclude with an open problem: Is it possible to design a compact distance oracle for a planar digraph which can handle multiple failures?

### References


Appendix

Avoiding the passing of blue edges as red edge of type 2

First we describe the situation under which a blue edge in \( G^r(\nu) \) could be passed as a red edge of type 2 to some l-child of \( \nu \). Consider an edge \( (u, v) \) in \( G^r(\nu) \) which is either a red edge of type 2 or a blue edge. Notice that both \( u \) and \( v \) lie outside \( \text{FAMILY}(\nu) \). The construction of \( G^r(\nu) \) from \( G^r(\mu) \) implies that there exists at least one vertex, say \( x \), in \( \nu.L \) such that \( (u, v) \) belongs to \( T^\mu(x); \) let \( w \) be the first such vertex from \( \nu.L \). If \( w = \text{FIRST}(\nu) \), then \((u, v)\) is a blue edge, otherwise it will be a red edge of type 2. In the latter case, it means that \((u, v)\) gets relaxed in the shortest path tree when the source vertex moves to \( w \) while executing MSSP in graph \( G^r(\nu) \). The graph \( G^r(\nu) \) is later transformed into \( G^r(\nu) \). Notice that difference between \( G^r(\nu) \) and \( G^r(\mu) \) is that the vertices of \( \nu.R \) can be used by any suffix avoiding shortest path. This can prepone or postpone the relaxation of edge \((u, v)\) while executing MSSP on \( G^r(\nu) \) as follows: when the source reaches vertex \( w \), the shortest path to \( v \) could pass through some vertices of \( \nu.R \) (see Figure 11). However, as the source further moves down \( \nu.L \), on reaching some vertex \( w' \), a shorter path to \( v \) is available which uses the edge \((u, v)\). It will be at this point that the edge \((u, v)\) would get relaxed. So the relaxation of edge \((u, v)\) gets postponed (in a similar fashion, it can be preponed). If \((u, v)\) was a blue edge in \( G^r(\nu) \), then this postponing of relaxation of \((u, v)\) will enforce it to be passed on as a red edge of type 2 to some of the l-children of \( \nu \). However, this can be avoided altogether as discussed below.

![Figure 11: Relaxation of edge \((u, v)\) gets postponed due to presence of \( \nu.R \) in \( G^r(\nu) \).](image)

Though MSSP on \( G^r(\nu) \) solves SPASS for destination vertices lying in \( G^r(\nu) \), actually we need to solve SPASS for destination vertices from \( \text{FAMILY}(\nu) \) only. This key observation motivates us to take the following simple-minded approach. For each blue edge or red edge of type 2 whose relaxation got preponed or postponed in \( G^r(\nu) \), we restore its relaxation index back to the index as specified by \( G^r(\mu) \) (see Subsection 5.1). As a result, a blue edge at \( \nu \) will be passed only as a blue edge to its l-children in \( \mathcal{H} \). Interestingly this restoration of the relaxation indices does not affect \( \nu.\text{SPASS}(w, x) \) for any \( w \in \text{FAMILY}(\nu) \) due to the following reason. If \( \nu.\text{SPASS}(w, x) \) does not pass through any vertex of \( \nu.R \), then \( \nu.\text{SPASS}(w, x) \) is the same as \( \mu.\text{SPASS}(w, x) \). In this case, the restoration of relaxation index of any edge does not affect \( \nu.\text{SPASS}(w, x) \). If \( \nu.\text{SPASS}(w, x) \) indeed passes through some vertex of \( \nu.R \), then \( \nu.\text{SPASS}(w, x) \) will not even pass through any blue or red edge of type 2 (see Figure 8).

All-pairs replacement paths in planar digraphs

Lipton and Tarjan [14] showed that every planar graph \( G \) on \( n \) vertices has a set of \( O(\sqrt{n}) \) vertices whose removal splits the graph into connected components each of maximum size \( 2n/3 \). Such a set \( S \) is called a balanced separator of \( G \). This result can be exploited to construct a tree data structure \( T(G) \) as follows. First compute a balanced separator \( S \) of \( G \). Let \( G_1, ..., G_k \) be the connected components (ignoring direction of the edges) of \( G \setminus S \). The root node of \( T(G) \) stores \( G \) and \( S \), and its children are \( T(G_i) \)’s built recursively for each \( 1 \leq i \leq k \). For any node \( \alpha \) in \( T(G) \), let \( G_\alpha \) be \( S_\alpha \) respectively the subgraph and the separator stored in it. It follows from the construction of \( T(G) \) that for each vertex \( v \in V \), there is a unique node \( \alpha \in T(G) \) such that \( v \in S_\alpha \). Also observe that \( T(G) \) has \( O(\log n) \) height. The data structure for all-pairs replacement paths problem is obtained by a suitable augmentation of \( T(G) \) as follows.

For each vertex \( s \in S_\alpha \), we build a data structure \( D(s) \) for the single source replacement paths in \( G_\alpha \), and also a similar data structure, \( D^R(s) \), in the graph obtained by reversing the edge directions in \( G_\alpha \). Notice the following key observation about this data structure. For any \( u, v, x \in V \), there exists a unique node \( \alpha \) in \( T(G) \) such that \( P(u, v, x) \) is present in \( G_\alpha \) and it passes through some vertex \( s \in S_\alpha \). Hence the data structures \( D(s) \) and \( D^R(s) \) together store \( P(u, v, x) \). In particular, \( \delta(s, v, x) + \delta^R(s, u, x) = \delta(u, v, x) \). We can use this observation to answer the query of \( \delta(u, v, x) \) as follows.

For all nodes \( \beta \) from \( \operatorname{LCA}(u, v) \) to the root in \( T(G) \), and for each \( s \in S_\beta \), we query \( D(s) \) and \( D^R(s) \) to compute \( \delta(s, v, x) + \delta^R(s, u, x) \), and report the minimum. The total number of such queries will be \( O(\sqrt{n}) \), and each query takes \( O(\log n) \) time. Thus it takes \( O(\sqrt{n} \log n) \) time to compute \( \delta(u, v, x) \).

The space occupied by the data structure described above is \( O(n^{3/2} \log^4 n) \). This can be easily generalized to achieve \( O(S) \) space and \( O(n^2/S \ \text{polylog} \ n) \) query time for any \( S \in [n^{3/2}, n^2] \) using \( \epsilon \)-division of planar graphs designed by Frederickson [9].