Detection of Spatially-Close Fiber Segments in Optical Networks

Muhammad Iqbal, Farabi; Trajanovski, Stojan; Kuipers, Fernando

DOI
10.1109/DRCN.2016.7470840

Publication date
2016

Document Version
Accepted author manuscript

Published in
2016 12th International Conference on the Design of Reliable Communication Networks (DRCN)

Detection of Spatially-Close Fiber Segments in Optical Networks

Farabi Iqbal, Stojan Trajanovski and Fernando Kuipers
Network Architectures and Services, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands
{M.A.F.Iqbal, S.Trajanovski, F.A.Kuipers}@tudelft.nl

***Abstract*—Spatially-close network fibers have a significant chance of failing simultaneously in the event of man-made or natural disasters within their geographic area. Network operators are interested in the proper detection and grouping of any existing spatially-close fiber segments, to avoid service disruptions due to simultaneous fiber failures. Moreover, spatially-close fibers can further be differentiated by computing the intervals over which they are spatially close. In this paper, we propose (1) polynomial-time algorithms for detecting all the spatially-close fiber segments of different fibers, (2) a polynomial-time algorithm for finding the spatially-close intervals of a fiber to a set of other fibers, and (3) a fast exact algorithm for grouping spatially-close fibers using the minimum number of distinct risk groups. All of our algorithms have a fast running time when simulated on three real-world network topologies.**

I. Introduction

Network services rely upon network fibers to communicate between network points-of-presence (PoPs). These fibers carry Terabits of data over long distances. Hence, various measures have been taken to ensure the fibers' robustness, e.g., coating against lateral forces, moisture and mechanical stress. However, fibers can and do still fail due to man-made disasters (e.g., construction backhoes, ships' anchor drops and terrorist attacks) or natural disasters (e.g., earthquakes, floods and volcanic eruptions). Fiber failures degrade and interrupt network services in the absence of proper service compensating measures, possibly leading to monetary penalties to network operators due to breached service level agreements.

Spatially-close fibers typically have a higher chance of failing simultaneously, due to a disaster, than fibers that are more distant. For instance, in 2006, an earthquake off southern Taiwan had cut eight undersea fibers in sixteen places, disabling offshore connectivity for China and Southeast Asia [1]. The repair time took months. In 2009, an earthquake off East Taiwan and Typhoon Morakot had cut eight undersea fibers [2]. Taiwan is known to be geographically positioned in the circum-Pacific seismic belt. Proper risk analyses on spatially-close fibers are thus crucial for providing survivable network services across similar risky geographical areas.

Fibers can be spatially close due to a variety of reasons. Fiber segments of different fibers may have been deployed in a single duct (a physical pipe for placing fibers between two locations) due to the high cost of digging ducts. Duct sharing between network operators is also beginning to become a norm [3], [4]. Duct sharing leads to overlapping fiber segments, as shown in Fig. 1. If the duct is cut, all the fibers within it will fail simultaneously. Fibers originating or terminating from/to a PoP may have spatially-close endpoints due to existing infrastructures or landscapes. For instance, if the PoP is next to a river, the fiber segments may have been placed closely alongside a bridge. Fibers in different ducts may also be spatially close due to the close proximity of their ducts.

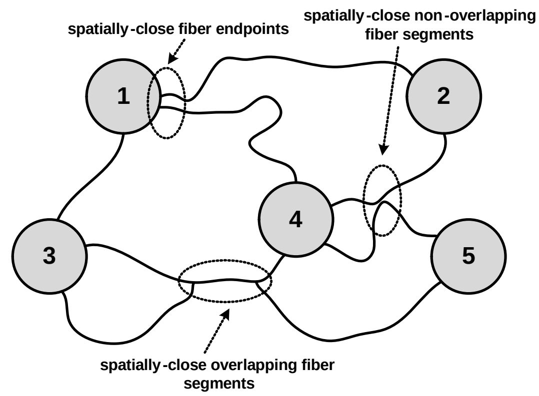
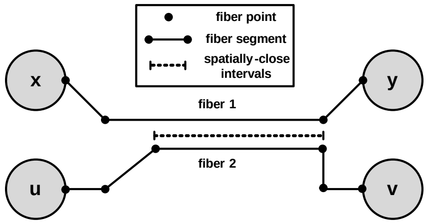


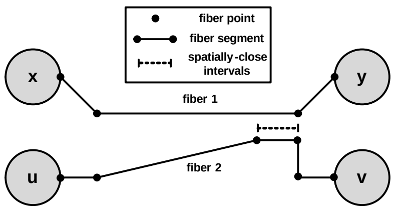
Fig. 1. Example of spatially-close fiber segments.

Confining to detecting only whether different fibers are spatially close can be misleading in terms of their risk of simultaneous failures. Fibers with longer spatially-close intervals have a higher chance of failing simultaneously compared to fibers with shorter spatially-close intervals, as shown by Fig. 2. For instance, there is a bottleneck of more than eight spatially-close fibers along the intervals of the Gulf of Suez [5]. Any disaster occurring along these intervals has a high chance of cutting all the fibers simultaneously, possibly leading to connectivity interruptions in many countries.

Our key contributions are organized as follows: (1) We propose fast polynomial-time algorithms in Section II for detecting all the spatially-close fiber segments of different fibers. (2) We propose a fast polynomial-time algorithm in Section III for detecting the intervals of a fiber that are spatially close to another fiber, and extend it for detecting the intervals of a fiber that are spatially close to a set of fibers. (3) We propose a fast exact algorithm in Section IV for grouping spatially-close fibers using a minimum number of risk groups. We simulate our algorithms on three real-world network topologies. We discuss related work in Section V and conclude the paper in Section VI.



(a)



(b)

Fig. 2. Fibers with different spatially-close intervals.

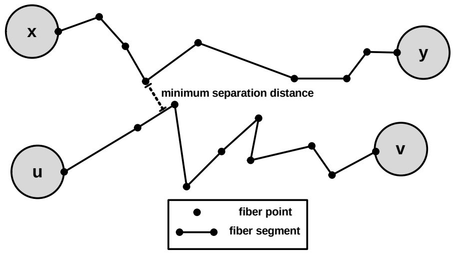


Fig. 3. Example of two fibers represented by our fiber structure.

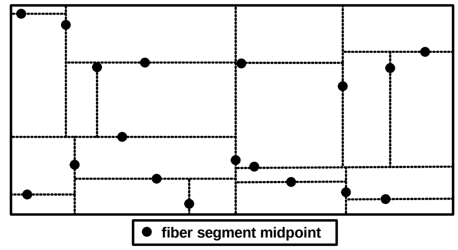
II. Detection of Spatially-Close Fiber Segments

A. Fiber Structure

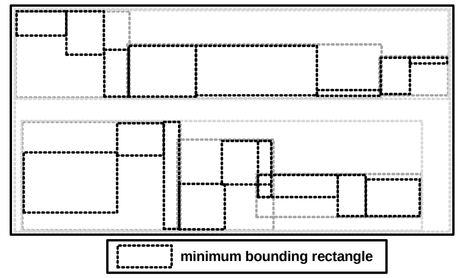
Fibers are commonly deployed in a non-straight manner, since deploying fibers in a straight line between PoPs can be cumbersome and impractical due to various reasons, e.g., terrain, existing infrastructure and government rules. Hence, we approximate fibers as non-straight concatenations of multiple fiber segments of irregular lengths, as shown in Fig. 3. The accuracy of the approximation increases with more fine-grained fiber segments. Each fiber segment is a straight line connecting two fiber points of known geodetic coordinates (latitude and longitude). For easier geospatial calculations, we assume that the geodetic coordinates of all fiber points can be projected into corresponding two-dimensional Cartesian coordinates. Two consecutive fiber segments may not necessarily have collinear fiber points (three or more points are collinear if they lie on a single straight line).

B. Problem Definition

Detection of Spatially-Close Fiber Segments (DSCFS) Problem: Given a set $\mathcal{L}$ of L fibers and a distance α. Each fiber $l \in \mathcal{L}$ is associated with a set $\mathcal{T}_l$ of T_l fiber segments. Each fiber segment $t \in \mathcal{T}_l$ is associated with two fiber points (u_{t_1}, v_{t_1}) and (u_{t_2}, v_{t_2}) of known geodetic locations. Find all the fiber segment pairs of different fibers that have a minimum separation distance of at most α.



(a) k-d tree partitioning



(b) R tree partitioning

Fig. 4. Example of fiber partitioning for tree representations.

The DSCFS problem is polynomially solvable (in the number of fiber segments) when the fiber structure of Section II-A is considered. One naive approach of solving the DSCFS problem is by computing the minimum separation distance between all fiber segment pairs of different fibers. If the minimum separation distance between any two fiber segments is at most α, they are spatially close. The worst-case time complexity of this naive approach is $O(L^2T^2)$, where T is the maximum number of fiber segments per fiber.

C. Our Approach

To achieve a faster practical running time, we propose Algorithm 1 that uses a k-d tree [6] to preprocess the fiber segments, and Algorithm 2 that uses an R tree [7] to preprocess the fiber segments for solving the DSCFS problem. A k-d tree is a space-partitioning data structure for organizing points in a k-dimensional space. An R tree is a depth-balanced data structure for organizing objects using bounded rectangles. The trees eliminate the need for computing the minimum separation distance between all fiber segment pairs, thus reducing the running time significantly.

Lines 1-3 of Algorithm 1 find the midpoint of all fiber segments and insert them into the k-d tree, since a k-d tree works with points instead of segments. Line 3 of Algorithm 2 computes the minimum bounding rectangle (MBR) of all fiber segments, and inserts them into the R tree. The MBR of a fiber segment is the smallest rectangle that encloses the fiber segment in the two-dimensional Cartesian plane. For instance, the partitioning of fibers of Fig. 3 by the k-d tree is shown in Fig. 4a, or by the R tree is shown in Fig. 4b. For each fiber segment t, lines 4-5 in Algorithm 1 find the set Z of fiber segments in Q with midpoints of a distance of at most $\alpha + \beta$ from the midpoint of fiber segment t, where β is the maximum fiber segment length. Similarly, for each fiber segment t, lines

Algorithm 1 DSCFS with k-d Tree Preprocessing

1: **for** each fiber $l \in \mathcal{L}$
2:     **for** each fiber segment $t \in \mathcal{T}_l$
3:         compute the segment midpoint m_t, and insert it into the k-d tree Q
4: **for** each fiber segment midpoint $m_t \in Q$ (m_t is the midpoint of fiber segment t)
5:     find the set Z of the entries of Q at a distance of at most $\alpha + \beta$ from m_t
6:     **for** each fiber segment midpoint $z_y \in Z$ (z_y is the midpoint of fiber segment y)
7:         **if** the minimum separation distance between fiber segments t and y is at most α
8:             fiber segments t and y are spatially close

Algorithm 2 DSCFS with R Tree Preprocessing

1: **for** each fiber $l \in \mathcal{L}$
2:     **for** each fiber segment $t \in \mathcal{T}_l$
3:         compute the segment MBR b_t, and insert it into the R tree Q
4: **for** each fiber segment MBR $b_t \in Q$ (b_t is the MBR of fiber segment t)
5:     find the set Z of the entries of Q at a distance of at most α from b_t
6:     **for** each fiber segment MBR $w_k \in Z$ (w_k is the MBR of fiber segment k)
7:         **if** the minimum separation distance between fiber segments t and k is at most α
8:             fiber segments t and k are spatially close

4-5 in Algorithm 2 find the set Z of fiber segment MBRs in Q with distance of at most α from the MBR of fiber segment t. Finally, lines 6-8 of Algorithm 1 and Algorithm 2 detect all the fiber segments of set Z that are spatially close (with minimum separation distance of at most α) to fiber segment t. Two scenarios need to be considered when computing the minimum separation distance between any two fiber segments:

Case 1: For intersecting fiber segments, their minimum separation distance is zero.

Case 2: Else, their minimum separation distance is the minimum distance between any points of the two fiber segments.

The ball tree [8] or newer variants of the R tree [9] can also be used instead with Algorithms 1 and 2. Although the worst-case time complexity of Algorithms 1 and 2 is similar to the naive approach (in the worst-case, all the fiber segments could be spatially close), significant time-savings can be expected in practice in solving the DSCFS problem.

D. Proof-of-Concept

We generate three real-world network topologies, namely the Angola Telecom network, the Ethiopia Telecom network and the Telkom South Africa network, from the Keyhole Markup Language (KML) datasets provided in [10]. We assume each placemark in the KML datasets as a distinct fiber. KML is a language schema of the Open Geospatial Consortium (OGC) for representing geographical information. We use this format in our proof-of-concept since KML datasets can be directly translated to our fiber structure. We use the projection wizard of [11] to generate an equidistant map projection for each network, such that all the two-dimensional plane coordinates (projected using [12]) have proportionally correct distances from the center point of the network geographic bounds. Of all the studied networks, the Telkom South Africa network has the highest number of fibers and fiber segments, the Angola Telecom network has the highest average length of fibers and fiber segments, while the Ethiopia Telecom network has the highest average number of fiber segments per fiber. Detailed properties of the networks are shown in Table I.

Algorithm 1 and Algorithm 2 are coded in Python and simulations were conducted on an Intel(R) Core i7-4600U 2.1GHz machine of 16GB RAM memory. Parsing the datasets and generating the network took less than a second for all the networks. As shown in Table II, the running time of the naive approach is quite high, even when the minimum separation distance α is low. Significant time-savings can be achieved by Algorithm 1 and Algorithm 2, particularly when α is low. Algorithm 2 outperforms Algorithm 1 since the performance of Algorithm 1 also relies on β. Although the running times of Algorithm 1 and 2 increase with α, α is often a small value. The running time of the algorithms is highest on the Telkom South Africa network, due to its high number of fibers and fiber segments.

III. Intervals of Spatially-Close Fibers

In the previous section, we discussed how to detect spatially-close fibers, and pinpoint spatially-close fiber segments. However, viewing spatially-close fibers as a Boolean relation can be misleading in terms of the risk of simultaneous failures. In this section, we find the intervals of spatially-close fibers, to give a better measure of the fibers' spatial proximity. For instance, the fiber pair in Fig. 2a has longer intervals that are spatially close than the fiber pair in Fig. 2b.

A. Intervals of A Pair of Spatially-Close Fiber Segments

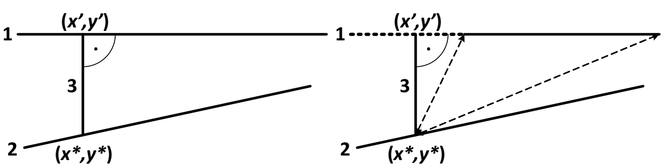
We first explain our approach of finding the intervals of a pair of spatially-close fiber segments, before proceeding with the intervals of spatially-close fibers in subsequent sections.

TABLE I
PROPERTIES OF THE STUDIED NETWORKS.

Property	Angola Telecom	Ethiopia Telecom	Telkom South Africa
Number of fibers	16	21	343
Total length of fibers	10943.86 km	8162.47 km	27849.70 km
Average length of fibers	683.99 km	388.69 km	81.19 km
Maximum length of fibers	1745.93 km	2105.52 km	479.08 km
Minimum length of fibers	261.53 km	67.35 km	4.01 km
Number of fiber segments	979	2917	4901
Average number of fiber segments per fiber	61.19	138.90	14.29
Maximum number of fiber segments per fiber	238	492	80
Minimum number of fiber segments per fiber	17	10	1
Average length of fiber segments	11.18 km	2.80 km	5.68 km
Maximum length of fiber segments	59.49 km	54.30 km	97.75 km

TABLE II
THE EFFECT OF α ON THE TIME TAKEN FOR SOLVING THE DSCFS PROBLEM.

Network	Algorithm	Minimum separation distance (α)						
		5 m	50 m	500 m	5 km	50 km	500 km	5000 km
Angola Telecom	Naive approach	20.16 s	20.16 s	20.16 s	20.16 s	20.16 s	20.16 s	20.16 s
	Algorithm 1	1.12 s	1.13 s	1.15 s	1.15 s	1.66 s	11.90 s	19.81 s
	Algorithm 2	0.21 s	0.22 s	0.23 s	0.23 s	0.34 s	9.17 s	21.20 s
Ethiopia Telecom	Naive approach	2.98 min	2.98 min	2.98 min	2.98 min	2.98 min	2.98 min	2.98 min
	Algorithm 1	7.12 s	7.37 s	7.55 s	7.80 s	14.79 s	2.12 min	3.05 min
	Algorithm 2	0.75 s	0.75 s	0.75 s	0.76 s	2.50 s	2.01 min	3.22 min
Telkom South Africa	Naive approach	8.60 min	8.60 min	8.60 min	8.60 min	8.60 min	8.60 min	8.60 min
	Algorithm 1	30.89 s	31.17 s	31.90 s	31.97 s	50.51 s	5.20 min	8.97 min
	Algorithm 2	1.28 s	1.30 s	1.31 s	1.45 s	8.23 s	4.53 min	9.00 min



(a) Case 1 (b) Case 2

Fig. 5. Intervals of a pair of spatially-close fiber segments

Suppose we have two fiber segments $y_1 = k_1x_1 + n_1$ of $x_1 \in [a,b]$ and $y_2 = k_2x_2+n_2$ of $x_2 \in [c,d]$. Let us fix a point $(x^*, y^* = k_2x^* + n_2)$ on fiber segment 2, as shown in Fig. 5. The y-intercept m of a line 3 through (x^*, y^*) that is perpendicular to fiber segment 1 is $m = (k_2 + \frac{1}{k_1})x^* + n_2$. The intersection of line 3 and the line projection of fiber segment 1 defines point (x', y')

$$-\frac{1}{k_1}x' + m = k_1x' + n_1$$
$$x' = \frac{1}{1+k_1^2}\left[(k_1k_2+1)x^* + k_1(n_2-n_1)\right] \quad (1)$$

For simplicity, consider $A = \frac{k_1k_2+1}{1+k_1^2}$ and $B = \frac{k_1(n_2-n_1)}{1+k_1^2}$. For the distance from all possible (x^*, y^*) to the projected line of fiber segment 1 to be at most α, the condition is

$$\begin{aligned}\alpha^2 &\geq (x'-x^*)^2 + (y'-y^*)^2 \\ &= ((A-1)x^* + B)^2 + ((k_1A-k_2)x^* \\ &\quad + k_1B + n_1 - n_2)^2 \end{aligned} \quad (2)$$

From Eq. 2, we have a quadratic inequality, whose solution is an interval $x^* \in [i_1, i_2]$. The intersection of this interval and the interval $[c,d]$ gives a necessary condition for x^*

$$x^* \in [z_1, z_2] = [i_1, i_2] \cap [c, d] \quad (3)$$

When (x', y') is on fiber segment 1, i.e. $x' \in [a,b]$

$$a \leq Ax^* + B \leq b \quad (4)$$

By solving Eq. 4 for x^*, we obtain an interval $x^* \in [j_1, j_2]$. The solution is then

$$x^* \in [z_1, z_2] \cap [j_1, j_2] \quad (5)$$

On the other hand, when (x', y') is not on fiber segment 1, i.e. $x' \notin [a,b]$, it is sufficient that the minimum of the distances from (x^*, y^*) to the two endpoints of fiber segment 1 is not bigger than α. In this case, $x^* \notin [j_1, j_2]$.

$$(x^*-a)^2 + (k_2x^* + n_2 - k_1a - n_1)^2 \leq \alpha^2, \text{ or} \quad (6)$$
$$(x^*-b)^2 + (k_2x^* + n_2 - k_1b - n_1)^2 \leq \alpha^2 \quad (7)$$

These two inequalities give two intervals $[p_1, p_2]$ and $[q_1, q_2]$. The solution is then

$$x^* \in [[z_1, z_2] \setminus [j_1, j_2]] \cap [[p_1, p_2] \cup [q_1, q_2]] \tag{8}$$

The solution gives the intervals of fiber segment 2 that are spatially close to fiber segment 1.

B. Problem Definition

Intervals of a Pair of Spatially-Close Fibers (IPSCF) Problem: Given two fibers l_i and l_j, and a distance α. Each fiber l_i/l_j is associated with a set $\mathcal{T}_i/\mathcal{T}_j$ of T_i/T_j fiber segments, respectively. Each fiber segment $t \in \mathcal{T}_i/\mathcal{T}_j$ is associated with two fiber points (u_{t_1}, v_{t_1}) and (u_{t_2}, v_{t_2}) of known geodetic locations. Find the intervals of fiber l_i that have a minimum separation distance of at most α to fiber l_j.

The IPSCF problem is an extension of the DSCFS problem. If two fibers are spatially close, detecting their spatially-close intervals is useful for recognizing to which extent they are vulnerable to simultaneous failures. The IPSCF problem is unidirectional, since the intervals of fiber l_i that are spatially close to fiber l_j are not necessarily equal to the intervals of fiber l_j that are spatially close to fiber l_i. When the network operator needs to find the spatially-close intervals of a fiber to a set of other fibers, the following problem is more useful.

Intervals to a Set of Spatially-Close Fibers (ISSCF) Problem: Given a fiber l_i, a set $\mathcal{Y}$ of Y fibers and a distance α. Each fiber l_i or $l_j \in \mathcal{Y}$ is associated with a set $\mathcal{T}_i/\mathcal{T}_j$ of T_i/T_j fiber segments, respectively. Each fiber segment $t \in \mathcal{T}_i/\mathcal{T}_j$ is associated with two fiber points (u_{t_1}, v_{t_1}) and (u_{t_2}, v_{t_2}) of known geodetic locations. Find the intervals of fiber l_i that have a minimum separation distance of at most α to any fiber $l_j \in \mathcal{Y}$.

The IPSCF problem is a subset of the ISSCF problem, where Y is a single fiber l_j. The IPSCF and ISSCF problems are polynomially solvable in the number of fiber segments.

C. Our Approach

We propose Algorithm 3 for solving the IPSCF and ISSCF problems. Lines 1-3 insert all the fiber segment MBR of each fiber $l_j \in \mathcal{Y}$ into an R tree Q. For each fiber segment t of fiber l_i, line 7 finds the set Z of fiber segment MBRs in Q with distance of at most α from the MBR of fiber segment t. Line 9 then uses the equations of Section III-A to find the spatially-close intervals of fiber segment t to each fiber segment $k \in Z$. Line 10 returns the intervals of fiber l_i that have a minimum separation distance of at most α from any fiber $l_j \in \mathcal{Y}$ by uniting all the intervals acquired by line 9. The worst-case time complexity of Algorithm 3 is YT^2, where T is the maximum number of fiber segments per fiber.

D. Proof-of-Concept

Algorithm 3 is coded in Python and simulations were conducted on an Intel(R) Core i7-4600U 2.1GHz machine of 16GB RAM memory. All simulation results are averaged over five thousand runs, and we randomly chose the fibers for l_i and $\mathcal{Y}$ in each simulation run. Fig. 6 shows the running time

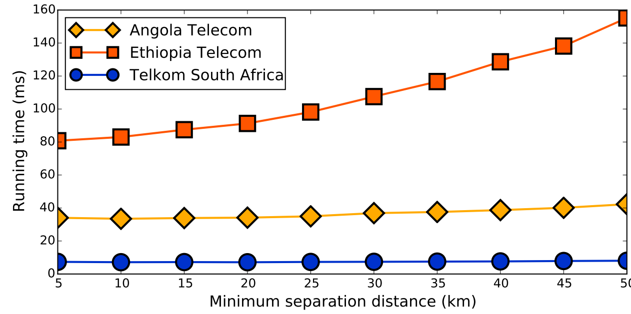


Fig. 6. Effect of α on the time taken to solve the ISSCF problem ($Y = 5$).

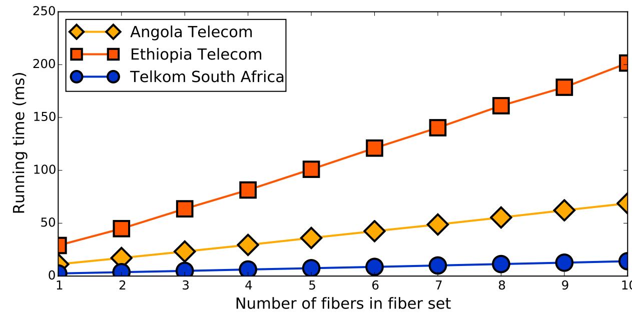


Fig. 7. Effect of Y on the time taken to solve the ISSCF problem ($\alpha = 25$ km).

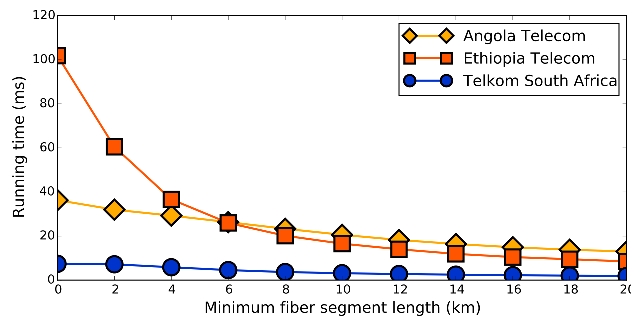


Fig. 8. Effect of limiting the length of fiber segment on the time taken to solve the ISSCF problem ($\alpha = 25$ km, $Y = 5$).

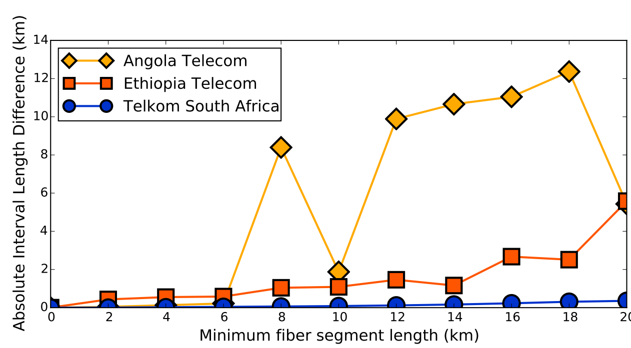


Fig. 9. Effect of limiting the length of fiber segment on the interval length of the ISSCF problem ($\alpha = 25$ km, $Y = 5$).

Algorithm 3 IPSCF with R Tree Preprocessing

for each fiber $l_j \in \mathcal{Y}$
  for each fiber segment $k \in \mathcal{T}_j$
    compute the fiber segment MBR b_k, and insert it into the R tree Q
set I as an empty interval
for each fiber segment $t \in \mathcal{T}_i$
  compute the fiber segment MBR b_t
  find the set Z of the entries of Q at distance at most α from b_t
  for each fiber segment MBR $w_k \in Z$ (w_k is the MBR of fiber segment k)
    $J \leftarrow$ the intervals of fiber segment t that have a minimum separation distance of at most α to fiber segment k
    $I \leftarrow I \cup J$

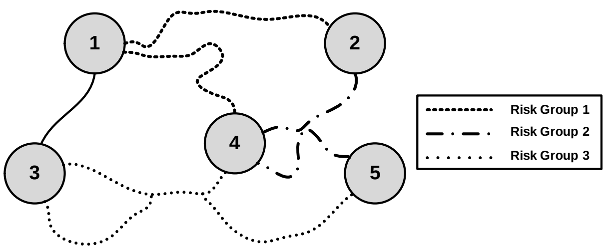


Fig. 10. An example of risk groups.

of Algorithm 3 as a function of α when Y is 5, while Fig. 7 shows the running time of Algorithm 3 as a function of Y when α is 25 km. Higher α or Y implies potentially longer intervals, such that the running time increases. The case when $Y = 1$ corresponds to the IPSCF problem.

To reduce the running time, fiber segments with collinear fiber points can be combined safely without affecting the results. Unfortunately, the number of collinear fiber segments in the studied datasets are too few for the effect to become apparent. It is also tempting to combine shorter fiber segments by imposing a lower bound on the fiber segment length, which reduces the number of fiber segments and the running time as shown in Fig. 8. However, by doing so, the network changes unpredictably, affecting the intervals as shown in Fig. 9.

IV. Grouping of Spatially-Close Fibers

This section builds upon Section II by grouping spatially-close fibers in the same risk group. A risk group is often defined as a set of fibers sharing an adjacent node or duct, e.g., [13], [14], [15]. We propose a broader definition of risk groups, such that a risk group is a set of fibers that are spatially close to every other fiber in the same set. Each risk group implies a set of fibers with a high chance of simultaneous failures. A risk group contains multiple fibers, and a fiber may belong to multiple risk groups. For instance, the fibers in Fig. 1 can be placed into three risk groups, as shown in Fig. 10.

A. Problem Definition

Grouping of Spatially-Close Fibers (GSCF) Problem: Given a set $\mathcal{F}$ of F spatially-close fiber pairs (each fiber pair has a minimum separation distance of at most α). Group all the fibers that are spatially close to each other, such that the number of distinct risk groups used is minimized.

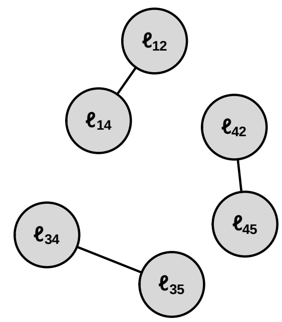


Fig. 11. Transformation between problems.

The GSCF problem is NP-hard due to the complexity of assigning the maximal risk groups. A maximal risk group is a set of fibers that are spatially close to every other fiber in the set, and which is not a subset of any other larger risk group. To prove that the GSCF problem is NP-hard, we show that any instance of the NP-hard Maximal Clique Enumeration (MCE) problem [16] can be transformed in polynomial time to an instance of the GSCF problem.

Maximal Clique Enumeration (MCE) Problem: Given a graph $H = (\mathcal{M}, \mathcal{E})$ of a set $\mathcal{M}$ of M nodes and a set $\mathcal{E}$ of E links. Each link $(u, v) \in \mathcal{E}$ connects nodes $u \in \mathcal{M}$ and $v \in \mathcal{M}$. Find all the maximal cliques in H. A maximal clique is a set of nodes that have a link to every other node in the set, and it is not a subset of any other larger clique.

We start our proof by adjusting the interpretation of H such that each node $m \in \mathcal{M}$ represents a distinct fiber $\ell \in \mathcal{F}$. H then consist of F nodes. If any two nodes $u \in \mathcal{M}$ and $v \in \mathcal{M}$ are connected by a link $(u, v) \in \mathcal{E}$ in H, the two fibers $\ell_u \in \mathcal{F}$ and $\ell_v \in \mathcal{F}$ are spatially close. For instance, the corresponding H for the fibers in Fig. 10 is shown in Fig. 11. By solving the GSCF problem in H, the MCE problem is solved as well, since each maximal risk group is equivalent to a maximal clique. Since the MCE problem (and thus the GSCF problem as well) is hard to approximate [17], we focus on a practically efficient exact approach to solve the GSCF problem.

Algorithm 4 GSCF

1: create an empty graph $H = (\mathcal{M}, \mathcal{E})$
2: **for** each fiber $l_i \in \mathcal{F}$
3:     **for** each fiber $l_j \in \mathcal{F} \setminus l_i$
4:         **if** fibers l_i and l_j are spatially close
5:             add node l_i and l_j in H
6:             connect node l_i and l_j by a link in H
7: find all the maximal cliques in H

B. Our Approach

We propose Algorithm 4, which is based on a graph transformation approach to solve the GSCF problem. An empty graph $H = (\mathcal{M}, \mathcal{E})$ is created in line 1. For each fiber pair in $\mathcal{F}$, a node is created for each of them in H in line 5, if a node representing them does not yet exist in H. Both nodes in H are connected by a link in line 6. The maximal risk groups can then be acquired by finding all the maximal cliques of graph H in line 7. Each resultant maximal clique represents a distinct maximal risk group. The maximal cliques can be acquired by the Bron-Kerbosch algorithm [18]. The variant of the Bron-Kerbosch algorithm by [19] has the best worst-case time complexity of $O(3^{\frac{M}{3}})$, since any graph with M nodes can have at most $3^{\frac{M}{3}}$ maximal cliques [20]. Using the implementation of [19], the worst-case complexity of Algorithm 4 is $O(3^{\frac{F}{3}})$ since there are at most F nodes in H.

C. Proof-of-Concept

Algorithm 4 is coded in Python and simulations were conducted on an Intel(R) Core i7-4600U 2.1GHz machine of 16GB RAM memory. We vary the minimum separation distance α (thus varying $\mathcal{F}$ via the DSCFS problem) while finding the maximal risk groups. The maximal risk group assignment takes a very short amount of time, less than a second in most of the tested cases as shown in Table III, with at most ten seconds for the Telkom South Africa network.

The maximum and minimum numbers of fibers per maximal risk group increase with the increase of α. Higher α increases the possibility of more fibers being identified as being spatially close to each other, and which consequently are assigned into the same maximal risk group. On the other hand, the total number of maximal risk groups can either increase or decrease with the increase of α. Higher α increases the possibility of more fibers being identified as being spatially close to each other, thus creating more risk groups, in which some of them are maximal. However, by having more fibers within each risk group, the possibility of a maximal risk group being a superset of another smaller risk group also increases, possibly reducing the number of maximal risk groups.

V. Related Work

The survivability of network services in the event of disasters has received increasing interest in recent years, e.g., [21], [22], [23], [24], [25], [26], [27]. The effect of disasters on network services can be studied from an attacker perspective, e.g., disconnecting two nodes under a disaster [22], how many disasters are needed to cut connectivity between different nodes [21] or the geographic area that brings the worse effect to the network connectivity when confronted by a disaster [23]. From an attacker perspective, the knowledge of the vulnerable network geographic areas (often modeled as a circular disk [21], an ellipse or a general polygon [23] or a half-plane [22]) is highly appealing. However, from a network operator perspective, knowing the vulnerable geographic areas without enough information on resolving the problem can be frustrating. There are endless possibilities of disaster shapes, and protecting against all of them is hard. In this paper, more attention is given to the fiber geodetic locations instead of the disaster shape. Spatially-close fibers have a high chance of failing simultaneously, regardless of the disaster shape.

We have proposed fast approaches on finding spatially-close fiber segments for any arbitrary positioning or length of fiber segments. While the work of [21], [23] may also be used out of context to solve a similar problem, their time complexity is much higher than ours, even without our practical time-saving preprocessing routines. [21] limits the fiber segments to be non-intersecting and two adjacent fiber segments are not collinear, while we do not impose such limitations. [23] is more suited for wireless networks instead of fiber networks, since they consider that only links adjacent to nodes inside a disaster area fail, while links merely passing through the area remain intact. Consider two perpendicular fibers with minimum separation distance from one of the fiber endpoints to the middle of the other (much longer) fiber. [23] would have not considered the fiber pair to be spatially close.

[26], [28] focus on a greenfield planning of placing new fibers in a geographic area. Our approach on solving the ISSCF problem is useful in complementing their work by verifying that the new fibers are not spatially close to any existing fibers.

[24] groups fibers by matching their locations to specific disaster maps. On the other hand, we focus on grouping fibers based on their spatial proximity. We also find the maximal risk groups, such that the number of risk groups needed are greatly reduced, which is favorable since many applications, e.g., risk-group-based routing, are NP-hard, thus the running time can grow exponentially with the number of risk groups.

VI. Conclusion and Future Work

We have proposed fast polynomial-time approaches for detecting spatially-close fiber segments. Network operators can adapt their existing network physical maps to our fiber structure for detecting spatially-close fiber segments in their network. We also showed that the algorithm based on R tree preprocessing far outperforms the algorithm with k-d tree preprocessing and the intuitive naive approach.

We have also proposed a polynomial-time approach for finding the intervals of a fiber that are spatially close to another fiber, i.e., the span of a fiber that is unsafe from the other fiber. These intervals can be used to differentiate spatially-close fibers according to their proximity. We also extended the problem to find the intervals of a fiber that are spatially

TABLE III
THE EFFECT OF α ON THE SIZES OF MAXIMAL RISK GROUPS.

Network	**Property**	**Minimum separation distance (α)**						
		5 m	**50 m**	**500 m**	**5 km**	**50 km**	**500 km**	**5000 km**
Angola Telecom	Total number of maximal risk groups	9	13	16	15	12	9	1
	Average number of fibers per maximal risk groups	2.22	2.15	2.25	2.60	3.00	8.11	16
	Maximum number of fibers per maximal risk groups	3	3	3	3	4	9	16
	Minimum number of fibers per maximal risk groups	2	2	2	2	2	7	16
Ethiopia Telecom	Total number of maximal risk groups	12	16	17	17	15	11	1
	Average number of fibers per maximal risk groups	2.08	2.25	2.41	2.41	2.67	11	21
	Maximum number of fibers per maximal risk groups	3	3	3	3	4	14	21
	Minimum number of fibers per maximal risk groups	2	2	2	2	2	10	21
Telkom South Africa	Total number of maximal risk groups	210	230	311	271	281	3659	1
	Average number of fibers per maximal risk groups	2.03	2.05	2.47	2.82	4.97	80.31	343
	Maximum number of fibers per maximal risk groups	3	3	5	8	9	108	343
	Minimum number of fibers per maximal risk groups	2	2	2	2	2	19	343

close to at least one fiber in a set of fibers. We also showed that maintaining the granularity of fiber segments is important, since combining non-collinear fiber segments changes the intervals unpredictably, albeit with significant time savings.

We have also proposed a fast exact approach for grouping spatially-close fibers using the minimum number of distinct risk groups. Our risk group classification enables ample knowledge of existing risk-group-disjoint-paths algorithms to be used in finding disaster disjoint-paths, and leads to a unified risk group classification when combined with existing risk group classification approaches. We showed that the number of maximal risk groups can increase or decrease with the increase of the minimum separation distance between fibers.

Examples of possible future work that can be derived from this paper are, 1) finding two paths P_1 and P_2 between two network nodes such that the minimum separation distance of the two paths is maximized, and 2) finding two paths P_1 and P_2 between two network nodes such that the total interval length of P_1 that is spatially-close to P_2 is minimized.

ACKNOWLEDGMENT

This research was supported by SURFnet.

REFERENCES

[1] http://www.telecomasia.net/content/subsea-cable-breaks-bad-2006-quake.

[2] http://www.submarinenetworks.com/news/cables-cut-by-taiwan-earthquake-and-typhoon-morakot.

[3] http://www.modernghana.com/news/414121/1/telcos-co-share-ducts-to-lay-fibre-cables.

[4] http://www.computing.co.uk/ctg/news/1937275/bt-reveals-hotly-anticipated-pricing-proposals-duct-pole-access.

[5] http://www.submarinecablemap.com.

[6] J. L. Bentley, "Multidimensional binary search trees used for associative searching," *Commun. ACM*, vol. 18, no. 9, pp. 509–517, 1975.

[7] A. Guttman, "R-trees: a dynamic index structure for spatial searching," in *ACM Special Interest Group Manage. Data (SIGMOD'84)*, 1984.

[8] S. M. Omohundro, "Five balltree construction algorithms," International Computer Science Institute Berkeley, Tech. Rep., 1989.

[9] L. Arge, M. De Berg, H. J. Haverkort, and K. Yi, "The priority R-tree: A practically efficient and worst-case optimal R-tree," in *ACM Special Interest Group Manage. Data (SIGMOD'04)*, 2004.

[10] https://github.com/stevesong/afterfibre-kml.

[11] http://projectionwizard.org.

[12] https://github.com/OSGeo/proj.4/wiki.

[13] J.-C. Bermond, D. Coudert, G. D'Angelo, and F. Z. Moataz, "Finding disjoint paths in networks with star shared risk link groups," *Elsevier Theor. Comput. Sci.*, vol. 579, pp. 74–87, 2015.

[14] P. Datta and A. K. Somani, "Graph transformation approaches for diverse routing in shared risk resource group (SRRG) failures," *Elsevier Comput. Netw.*, vol. 52, no. 12, pp. 2381–2394, 2008.

[15] S. Yang, S. Trajanovski, and F. Kuipers, "Availability-based path selection and network vulnerability assessment," *Wiley Networks*, 2015.

[16] E. A. Akkoyunlu, "The enumeration of maximal cliques of large graphs," *SIAM J. Comput.*, vol. 2, no. 1, pp. 1–6, 1973.

[17] M. R. Garey and D. S. Johnson, "Strong NP-completeness results: Motivation, examples, and implications," *J. ACM*, vol. 25, no. 3, pp. 499–508, 1978.

[18] C. Bron and J. Kerbosch, "Algorithm 457: finding all cliques of an undirected graph," *Commun. ACM*, vol. 16, no. 9, pp. 575–577, 1973.

[19] E. Tomita, A. Tanaka, and H. Takahashi, "The worst-case time complexity for generating all maximal cliques and computational experiments," *Elsevier Theoret. Comput. Sci.*, vol. 363, no. 1, pp. 28–42, 2006.

[20] J. W. Moon and L. Moser, "On cliques in graphs," *Springer Israel J. Math.*, vol. 3, no. 1, pp. 23–28, 1965.

[21] S. Neumayer, A. Efrat, and E. Modiano, "Geographic max-flow and min-cut under a circular disk failure model," *Elsevier Comput. Netw.*, vol. 77, pp. 117–127, 2015.

[22] H. Saito, "Analysis of geometric disaster evaluation model for physical networks," *IEEE/ACM Trans. Netw.*, 2014.

[23] S. Trajanovski, F. A. Kuipers, A. Ilic, J. Crowcroft, and P. Van Mieghem, "Finding critical regions and region-disjoint paths in a network," *IEEE/ACM Trans. Netw.*, vol. 23, no. 3, pp. 908–921, 2015.

[24] F. Dikbiyik, M. Tornatore, and B. Mukherjee, "Minimizing the risk from disaster failures in optical backbone networks," *IEEE J. Lightw. Technol.*, vol. 32, no. 18, pp. 3175–3183, 2014.

[25] C. Doerr and F. Kuipers, "All quiet on the internet front?" *IEEE Commun. Mag.*, vol. 52, no. 10, pp. 46–51, 2014.

[26] C. Cao, M. Zukerman, W. Wu, J. H. Manton, and B. Moran, "Survivable topology design of submarine networks," *IEEE J. Lightw. Technol.*, vol. 31, no. 5, pp. 715–730, 2013.

[27] F. A. Kuipers, "An overview of algorithms for network survivability," *ISRN Commun. Networking*, 2012.

[28] D. L. Msongaleli, F. Dikbiyik, M. Zukerman, and B. Mukherjee, "Disaster-aware submarine fiber-optic cable deployment," in *IEEE Int. Conf. Opt. Netw. Design Modeling (ONDM'15)*, 2015.