Route Selection for Cabling Considering Cost Minimization and Earthquake Survivability via a Semi-Supervised Probabilistic Model

Mingbo Zhao, *Member, IEEE,* Tommy W. S. Chow, *Senior Member, IEEE,* Peng Tang, Zengfu Wang, Jun Guo, *Member, IEEE,* and Moshe Zukerman, *Fellow, IEEE*

***Abstract*—This paper focuses on an important and fundamental problem of connecting two points by a cable, subject to a tradeoff between cost and earthquake survivability. In particular, we address the problem of selecting a route for laying a cable under arbitrary topography, based on earthquake data. First, we derive a semi-supervised probability density estimation model for the likelihood of earthquake disaster. Based on this probabilistic model, we generate a nearest neighbor graph. The graph represents each data point with a four-dimensional space formed by the 3-D undersea coordinates and the relevant earthquake hazard level. It then forms the weight on graph between any positions. The data used in this study are all real data of undersea topography and earthquake information of the Taiwan Strait. As a result, both the undersea topology and the earthquake level can be transferred into a distance for shortest route finding. Finally, Dijkstra's algorithm is used for finding the optimal shortest route for cabling between the two given points on the graph. Extensive simulations based on a synthetic dataset and the Taiwan Strait real-world dataset corroborate the effectiveness of the proposed method.**

***Index Terms*—Cable route planning, kernel density estimation, semi-supervised learning, telecommunication cabling.**

I. Introduction

Route planning is an important topic in many real-world applications such as mobile robot path planning [1], real-time unmanned aerial vehicle (UAV) path planning [2], network design [3] and cable laying [4]–[6]. The goal of route planning is to find an optimal route between two given locations with certain requirements satisfied. Optimal route finding has always been associated with the shortest route and deterministic search algorithms have been used to find the exactly shortest route [1], [7], [8]. But different applications require different metrics and considerations when searching for the shortest route. For example, Ganganath, *et al.* [1] considered uneven terrains as a key factor to find energy-efficient paths for a mobile robot. Roberge *et al.* [2] found the best realtime UAV path by considering various factors such as average altitude, fuel consumption and radar exposure. Similar considerations can be applied to the design of a route for laying undersea cable, where it is important to consider potential undersea hazards (such as shark bites, ship anchors and earthquakes) instead of considering only cabling length or cost. In this paper, we will focus on handling such an important problem faced by engineers and surveyors of how to plan a route at minimal cost considering not only the topography, but also future hazards from disasters such as earthquakes.

During the past decade, there has been a series of disasters that caused major disruptions to Internet services. These include the earthquake of 7.1-magnitude in Hengchun, Southern Taiwan on 26th December, 2006 that severely damaged seven submarine communication cables and resulted in a major disruption to the international communications in South East Asia [9], the multiple-cable breaks that occurred twice in the Mediterranean Sea in 2008 [10], and the massive earthquake of 9.0-magnitude followed by a tsunami occurred near the east cost of Hongshu, Japan on 11th March 2011 that destroyed at least 2126 roads and 26 railways [11]. It is likely that similar disasters will occur in the future. Given that the global business and economy are highly dependent upon the functioning of modern infrastructures, it is important to find ways to mediate the adverse effects of disasters to avoid grave financial consequence. Survivability of cables, roads and railways, especially when they are positioned in an earthquake prone region, is important and should be considered during the design stage. Thus, we only focus on Internet optical fiber cables, but our results and discussions apply also to other cables, roads and railways.

Recognizing the potential cost associated with damage to Internet cables, cable survivability has been considered an important research topic [4]–[6]. There are many causes for cable break including but not limited to shark bites, ship anchors and earthquakes. One approach to achieve cable survivability is to guarantee sufficient (but not excessive) redundant capacity so that traffic can be temporarily routed away from a damaged cable through other available cables [12]–[15]. Another approach is to analyze or design a network topology considering correlated failures due to the effect of a disaster [16]–[18]. However, these publications, which focused on network survivability and resilience, have not considered the fundamental problem of minimizing the likelihood of an individual cable break caused by an earthquake. Considering

Copyright ©2016 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

The work described in this paper was supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CityU8/CRF/13G).

M. Zhao, T. W. S. Chow, P. Tang, J. Guo, and M. Zukerman are with the Department of Electronic Engineering, City University of Hong Kong, Hong Kong (e-mail: mbzhao4@gmail.com; eetchow@cityu.edu.hk; rollegg@gmail.com; j.guo@cityu.edu.hk; m.zu@cityu.edu.hk).

Z. Wang is with the School of Automation, Northwestern Polytechnical University, Xi'an, China (e-mail: wangzengfu@gmail.com).

route planning of a single cable between two end-points, we focus here on optimizing a cost function of the Euclidean distance and the earthquake disaster level.

Unlike previous studies of disaster effects on telecommunications cabling that considered a network of cables and network survivability, we consider the fundamental problem of how to optimize the shape of a single cable when it is laid between two points and passing an active earthquake area over a three-dimensional landform. We note that there is a proprietary commercial product [19] for cable route planning, but it does not consider earthquake survivability and its methodology is unpublished. Cao [20] and Cao *et al.* [21] considered cable routing between two points, but they only considered a two-dimensional landform, so they did not consider the geography, and the analyses there were only performed for special cases of cable shape.

In this paper, we propose an approach for cable route selection by considering both cabling distance and earthquake risk factor, in which we model the cable route in the three-dimensional landform map and earthquake likelihood map. To form the earthquake likelihood map, we model the earthquake risk factor by using a semi-supervised kernel density estimation model [22]. The semi-supervised probabilistic model propagates the disaster levels of earthquake from labeled positions (areas with earthquake records) to unlabeled positions (areas without earthquake record). Thus, the likelihood of earthquake for unknown areas can be estimated. Using this probabilistic model, we construct a graph using the neighborhood information. The graph represents each data point with four-dimensional attributes formed by the three-dimensional undersea coordinates and the one-dimensional earthquake hazard map. It then forms the weight on graph with the distance between any nearby data point. The data used in this study are all real data of undersea topography and earthquake information of the Taiwan Strait. As a result, both the undersea topology and the earthquake disaster level are transferred into a distance for finding the shortest route. Finally, Dijkstra's algorithm, a well-known graph-based method, is used to search the optimal shortest route between any two locations on the graph. Extensive experiments based on a synthetic dataset and a real-world dataset are presented. The results demonstrate the effectiveness of the methodology and show that the cabling route can achieve a good compromise between minimizing the cabling cost and earthquake risk.

The main contributions of the paper are as follows. First, we propose a new approach for cable route planning. This approach enables us to consider the Euclidean data and the earthquake risk factors together. Second, we apply the semi-supervised kernel density estimation method of [22] to construct the probabilistic model of earthquake likelihood map. Third, we introduce a new concept of using a graph manifold to model the topology and level of risk caused by earthquake disaster. Fourth, by modeling the problem in the above way, we show that the entire cabling design problem can be effectively solved by using the Dijkstra's algorithm.

The rest of the paper is organized as follows. Section II describes and motivates the route selection problem. In Section III, we present the semi-supervised kernel density estimation model to predict the disaster level of earthquake, and design the graph manifold to capture the structure of topology and evaluate the risk level of disaster. Extensive experiments are given in Section IV to show the effectiveness of the proposed method, and the final conclusion is drawn in Section V.

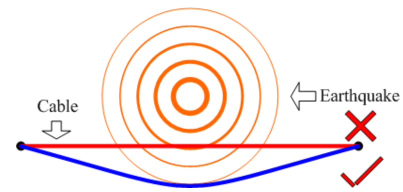


Fig. 1: Disaster survivability: the earthquake breaks the straight-line (red line); however, laying curved cable can avoid the consequence of earthquake (blue line).

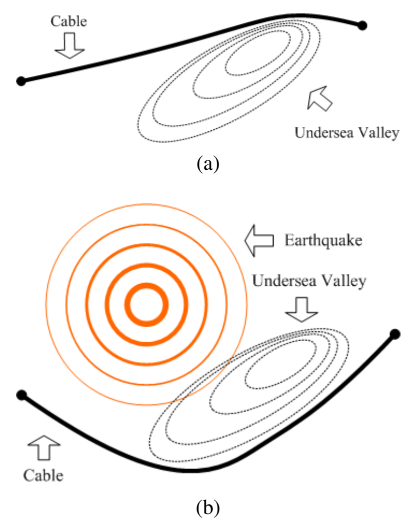


Fig. 2: An illustration of cost savings by avoiding a deep valley (the dotted contours) and earthquake (the orange circles): (a) only the landform is considered; (b) both the landform and earthquake risk are considered.

II. Problem Description

Normally, laying a cable in a straight line minimizes its length. However, for various reasons cables cannot be laid in straight lines. One important factor is the undersea topography which results in curve cabling. For example, deep valleys undersea and rough landforms have to be avoided due to their laying cost and cable-weight implications. Also, reducing the length of a cable in a risk prone area (e.g., earthquake faults) will reduce the risk of cable break. Fig. 1 shows a cable which is required to avoid the earthquake region. It is clear from Fig. 1 that, if this important issue of finding cable route to survive a disaster is not considered by cable designers, the consequence can be a significant disruption of communications services. Fig. 2 illustrates the impact of topography, as well as earthquake risk, on cable curving. Rough landforms may damage cables and therefore should be avoided. Meanwhile, seismically active areas usually accompany rough topography. Consequently, the impact of topography and earthquake risk on cable should be considered together.

Unlike previous studies, in which the topology was assumed to lie on a two-dimensional plane, we will model the surface of the Earth more accurately as a two-dimensional manifold in three-dimensional space. Different areas, such as land, sea, mountains or valleys, incur different cabling cost. The objective is to minimize the total cable cost that takes into account the cost per kilometer in the different sections of the region that we model, which can be solved using shortest route search methods such as Dijkstra's algorithm. However, using these methods to search the shortest route does not consider the disaster survival probability. This motivates us to design an effective approach of cost minimization that also satisfies an earthquake survivability requirement.

III. Methodology

In this paper, we focus our study on the fundamental problem of planning the route of a single cable on the Taiwan Strait. The public earthquake hazard map of this area, however, is not available and limited by the exploration and remote sensing technologies. In addition, the only accessible hazard information is the distinct earthquake records of the past 100 years. To utilize these seismic records for this study, we propose a probability density estimation method based on the available earthquake records to estimate the seismic risk of other areas of the Taiwan Strait. Specifically, with the assistance of the limited past 100-year earthquake records on the Taiwan Strait, we use Gaussian decay to estimate the risk in the open sea area where detailed seismic risk data is unavailable.

Here, we first introduce several key notations used in our work. Let the region be divided into n small square grids in the two-dimensional map formed by the longitude and latitude. Each grid node is represented by a position and is associated with d attributes including its topology information and earthquake information. Thus, we have a set of n positions, i.e., $\{x_1, \ldots, x_n\} \in R^{d\times n}$, in the region where the cable is planned to pass through. We define that a position is of class D if it is in an earthquake-prone area. Correspondingly, a position is said to be of class $\bar{D}$ if it is not in an earthquake-prone area. We denote $p(D|x)$ as the probability that a given position x is of class D, and $p(x|D)$ as the distribution of positions that are of class D. In practice, the likelihood of earthquake eruption can be estimated for some locations, based on information of previous earthquakes. Accordingly, let the first l positions be labeled, i.e., $p(D|x_j)$ is given for $1 \leq j \leq l$. Then, the remaining $n-l$ positions are unlabeled, i.e., $p(D|x_j)$ needs to be estimated for $l+1 \leq j \leq n$. For clarity, we list all the key notations and their descriptions in Table I.

TABLE I: Notations

Notation	Description
D	Class of positions that are in an earthquake-prone area
$\bar{D}$	Class of positions that are not in an earthquake-prone area
n	Total number of positions
d	Number of attributes associated with each position
l	Number of labeled positions
$p(D\|x)$	Probability that a given position x is of class D
$p(x\|D)$	Distribution of positions that are of class D
$p(D)$	Prior probability of class D
$k(x)$	Kernel function

A. Probability Density Estimation for Predicting Earthquake Disaster Level

It should be noted that earthquake survivability is an important factor for telecommunication cabling. Hence how to obtain the distribution of positions that are in an earthquake-prone area is a major concern for safety cabling design. There have been studies on highlighting areas according to their vulnerability to hazard, such as earthquake, called hazard mapping, to provide guidance for preventing serious damage [23]. Seismic hazard maps are used for planning and insurance purposes for large-scale projects, where cable systems, power transmission systems and railways pass through hazardous areas. However, such hazard maps, for example, those produced based on studies of peak ground motion, are very expensive, and only rich countries can afford them. For example, USGS provides seismic hazard maps with budget of over a billion dollar per year [19].

In this subsection, based on the work of [22], we present a probability density estimation model to approximate the hazard mapping of earthquake. Here, our goal is not to predict undersea earthquake but only to form an earthquake likelihood map, which is useful for designing a survivable cabling system. It is well known that the density estimation can be categorized into two different classes of approaches, parametric and non-parametric. Among all non-parametric approaches, kernel density estimation (or Parzen density estimation) is the most popular and has been widely used in many applications. Specifically, the distribution of positions that are of class D can be given as

$$p(x|D) = \frac{1}{l}\sum_{x_j \in D} k(x - x_j) \tag{1}$$

where $k(x)$ is a kernel function that satisfies $k(x) > 0$ and $\int k(x)dx = 1$. There are many ways to define the kernel function; a Gaussian kernel is one such choice.

It should be noted that $p(x|D)$ is the probability density function conditional on class D. However, when designing the route for cabling, we need to estimate the posterior probability $p(D|x_j)$ at each position x_j for $l+1 \leq j \leq n$. Let us first reformulate (1) in a weighted form as

$$p(x|D) = \frac{\sum_{j=1}^{n} p(D|x_j)k(x - x_j)}{\sum_{j=1}^{n} p(D|x_j)}. \tag{2}$$

Note that the kernel density function defined in (2) is equivalent to the one in (1) if $p(D|x_j) = 1$ for $1 \leq j \leq l$ and $p(D|x_j) = 0$ for $l+1 \leq j \leq n$. Hence (2) is an extension of (1).

We next show how to estimate $p(D|x_j)$ at each position x_j for $l+1 \leq j \leq n$. Let $p(D)$ be the prior probability of class D. Following the Bayes' rule, the posterior probability

$p(D|x_j)$ can be calculated as

$$p(D|x_j) = \frac{p(x_j|D)p(D)}{p(x_j|D)p(D) + p(x_j|\bar{D})p(\bar{D})}. \tag{3}$$

The prior probability $p(D)$ can be approximated as

$$p(D) \approx \frac{1}{n}\sum_{j=1}^{n} p(D|x_j). \tag{4}$$

Then, based on (2)-(4), we can derive the posterior probability $p(D|x_j)$ as

$$\begin{aligned} &p(D|x_j) \\ &= \frac{\sum_{i=1}^{n} p(D|x_i)k(x_j - x_i)}{\sum_{i=1}^{n} p(D|x_i)k(x_j - x_i) + \sum_{i=1}^{n} p(\bar{D}|x_i)k(x_j - x_i)} \\ &= \frac{\sum_{i=1}^{n} p(D|x_i)k(x_j - x_i)}{\sum_{i=1}^{n} k(x_j - x_i)}. \end{aligned} \tag{5}$$

By (5), we can obtain $p(D|x_j)$ at each position x_j for $l+1 \leq j \leq n$. For convenience, let $\pi_j = p(D|x_j)$. We write $\pi^{\mathrm{L}} = \{\pi_j\}$ for $1 \leq j \leq l$ and $\pi^{\mathrm{U}} = \{\pi_j\}$ for $l+1 \leq j \leq n$. Let $P = \{P_{ij}\}$ where $P_{ij} = k(x_j - x_i)/\sum_{i=1}^{n} k(x_j - x_i)$ for $1 \leq i, j \leq n$. Then, we split the matrix P into four blocks where $P^{\mathrm{LL}} = \{P_{ij}\}$ for $1 \leq i, j \leq l$, $P^{\mathrm{LU}} = \{P_{ij}\}$ for $1 \leq i \leq l$ and $l+1 \leq j \leq n$, $P^{\mathrm{UL}} = \{P_{ij}\}$ for $l+1 \leq i \leq n$ and $1 \leq j \leq l$, and $P^{\mathrm{UU}} = \{P_{ij}\}$ for $l+1 \leq i, j \leq n$. Based on (5), we can calculate π^{U} as

$$\pi^{\mathrm{U}} = \pi^{\mathrm{L}} P^{\mathrm{LU}} + \pi^{\mathrm{U}} P^{\mathrm{UU}} \Rightarrow \pi^{\mathrm{U}} = \pi^{\mathrm{L}} P^{\mathrm{LU}} \left(I - P^{\mathrm{UU}}\right)^{-1} \tag{6}$$

where I is the identity matrix. After we have estimated the disaster level of all positions using (6), we can incorporate them into the construction of graph by considering both the geometrical three-dimensional topology and disaster level of earthquake.

B. A Scalable Solution for the Probabilistic Model

It should be noted that the exact solution of (6) requires the computation of the inverse of $I - P^{\mathrm{UU}} \in R^{(n-l)\times(n-l)}$. This has a computation complexity of $\mathcal{O}((n-l)^3)$ which can be prohibitive for a large n. To reduce this huge computational cost, we adopt an iterative approach to calculate π^{U} in this subsection for handling large-scale dataset. Specifically, consider the matrix property

$$\left(I - P^{\mathrm{UU}}\right)^{-1} = \sum_{k=0}^{\infty} \left(P^{\mathrm{UU}}\right)^{k}. \tag{7}$$

Let us define

$$\pi^{\mathrm{U}}(t) = \pi^{\mathrm{L}} P^{\mathrm{LU}} \sum_{k=0}^{t} \left(P^{\mathrm{UU}}\right)^{k}. \tag{8}$$

It is clear that $\pi^{\mathrm{U}}(t)$ of the form (8) converges to π^{U} when t approaches ∞. From (8), we derive

$$\pi^{\mathrm{U}}(t) = \pi^{\mathrm{U}}(t-1) + \pi^{\mathrm{L}} P^{\mathrm{LU}} \left(P^{\mathrm{UU}}\right)^{t} \tag{9}$$

where we set $\pi^{\mathrm{U}}(0) = \pi^{\mathrm{L}} P^{\mathrm{LU}}$.

Compared with the exact solution of (6), using the iterative procedure of (9) requires a computation complexity of $\mathcal{O}((n-l)T)$, where T is the number of iterations. As can be seen in the simulation result of the next section, the iterative procedure achieves an accurate solution within only several hundred iterations, i.e., $T \ll (n-l)$. This means that the computational requirement of the iterative procedure is much smaller than that of the exact solution based on (6), which guarantees scalability to any realistic sized dataset. In all of the experimental study, we adopt the iterative procedure of (9) for estimating the earthquake disaster level instead of using the exact solution of (6).

C. Automatic Route Selection

The Dijkstra's algorithm is a widely used graph based method to find the shortest route between two nodes along a graph. The key step is to construct an undirected graph associated with weights on edges. Here, for the cable routing problem, the nodes on the graph represent the discrete three-dimensional coordinates of the topology, and the weights on edges represent certain measurements between nearby coordinates. Our goal is to develop a graph which can take the cabling cost and natural disaster (such as earthquake) into account when designing cable routing. As a result, the shortest route on the graph determined by the Dijkstra's algorithm finds effectively the optimal solution in terms of a tradeoff between the cabling cost and the cable disaster survivability.

Studies have shown that the cable bending angle caused by the rough landform plays an essential role on the survivability of the cable [18], which requires the route to be "smooth". We can observe that using only three-dimensional coordinates as the attributes of each position cannot sufficiently reflect the risk and disaster level of earthquake. Hence for any position x_j we define its attributes as $[X_j, Y_j, Z_j, E_j]$ where the first three elements are the three-dimensional X-Y-Z coordinates of position x_j in the topology, and

$$E_j \overset{\mathrm{def}}{=} \log p(D|x_j) \tag{10}$$

represents the earthquake disaster level at position x_j. Based on these attributes, we can design the weights on edges between any nearby positions. In particular, given two positions x_i and x_j, we define the (weighted) distance between them as

$$\begin{aligned} W(x_i, x_j) = &\sqrt{(X_i - X_j)^2 + (Y_i - Y_j)^2 + \gamma^2 (Z_i - Z_j)^2} \\ &+ \lambda(E_i + E_j)/2. \end{aligned} \tag{11}$$

In (11), the first term on the right-hand side represents the Euclidean distance between x_i and x_j in the three-dimensional coordinates of topology, and the last term represents a measure of seismic risk for the cable route to avoid risky areas, where γ and λ are parameters for balancing the tradeoff between the first term and the last term. From (11), we can see that $W(x_i, x_j)$ utilizes the average of E_i and E_j to evaluate the earthquake factor of the link between x_i and x_j. In addition, it is symmetrical, i.e., $W(x_i, x_j) = W(x_j, x_i)$, which can be directly used for the Dijkstra's algorithm.

After we form the four-dimensional attributes and design the weights on edges following (11), we can generate a nearest neighbor graph by connecting each node to its four neighborhoods in the X-Y coordinates. This is reasonable as the connection between a node and its neighbors are usually direct. We then search the shortest route along the graph for telecommunication cabling.

IV. Experimental Results

A. An Example based on Synthetic Dataset

We first evaluate the proposed method based on a synthetic data. The data are sampled from the region of $\{(X,Y)\,|X\in[-3,3],Y\in[-3,3]\}$, where a mountain is generalized around the point $[2.5,2.5]$ and an earthquake is generalized at $[-2.5,-2.5]$. In our study, we first use the probabilistic model in Section III-A to estimate the earthquake disaster level in order to design the route where the overall disaster level along the route can be minimized. We then use the proposed method in Section III-C to find the shortest route from the starting point $[3,3]$ to the end point $[-3,-3]$.

The simulation settings are as follows. Since the parameters γ and λ in (11) control the tradeoff between the earthquake disaster level and the cabling cost bypassing a mountain or valley, we consider four cases in our simulation in order to show the superiority of the proposed method. Specifically, the four cases are:

- $\lambda=0$ and $\gamma=0$. In this case, setting $\lambda=0$ and $\gamma=0$ means that the factors of earthquake and topology are not considered when searching the shortest route.
- $\lambda=100$ and $\gamma=0$. In this case, only the factor of earthquake is considered as λ is set to a large value.
- $\lambda=0$ and $\gamma=100$. In this case, only the factor of topology is considered.
- $\lambda=100$ and $\gamma=100$. In this case, the factors of earthquake and topology are all considered.

The simulation results are given in Fig. 3, where the first two columns show the shortest routes across the rough terrain in a three-dimensional topology and a two-dimensional contour surface, respectively, while the last column shows those across the earthquake area. The shortest route is denoted by a red line in all cases. From the simulation results in Fig. 3, we have the following observations. First, the route learned in Fig. 3(a) is the shortest. However, since the route learned by setting $\lambda=0$ and $\gamma=0$ will not consider the factors of terrain and earthquake, it crosses the earthquake area and rough terrain, i.e., the mountain peak. This is not practical in the real-world undersea cabling design. Second, the obtained results vary given different values of λ and γ. For example in Fig. 3(d), by carefully adjusting the parameter λ and γ, we observe that the shortest route can well bypass the earthquake area or rough terrain. Another observation is that the routes obtained in Fig. 3(d) are much longer than those in Fig. 3(a), which may increase the undersea cable cost. But in practice, we can carefully adjust the parameters in the proposed method to achieve a balancing result, where the cabling distance and earthquake risk factor can be optimized altogether.

B. Data Construction for Real-World Application

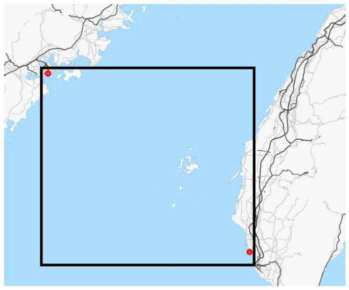
In order to evaluate the proposed method, we chose Taiwan Strait topographical data for our study. The data is obtained from the National Oceanographic Data Center (NODC: http://www.nodc.noaa.gov/access/index.html). The selected data are collected from an area 22.9N to 24.5N, 118.1E to 120.1E which covers 230 x 200 square kilometers. There are submarine mountains and canyons in the area. Moreover, the selected area locates on the Ring of Fire, an area where a large number of earthquakes and volcanic eruptions occurred in the basin of the Pacific Ocean. Undersea explorations have found that there are 1km-deep undersea canyons and severe terrain fluctuations. Figs. 4(a) and 4(b) show the area and the corresponding three-dimensional topology of under-sea floor, where a point with dark color on the map indicates a larger depth. From Fig. 4(b), we can observe a huge submarine canyon and rough landforms in the selected area. In addition, we have also collected 451 earthquake data at this seafloor occurred in the selected Taiwan Strait area from USGS Earthquake Hazards Program (http://earthquake.usgs.gov/). Fig. 4(c) shows such earthquake data with different colors for distinguishing overlapped earthquake centers. In Fig. 4(c), the size of a point is proportional to the corresponding earthquake magnitude. Our study is to find routes between two cities represented by two red points in Fig. 4(a), where the output routes should be as short as possible and at the same time can avoid the rough landform and earthquake areas.

C. Simulation Results for Real-World Application

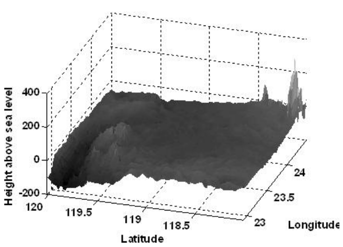
In this subsection, we evaluate the proposed method based on Taiwan Strait topographical data. Recall that in Section III-A we developed a probabilistic model for estimating the earthquake disaster level in order to design the route where the overall disaster level at the positions along the route can be minimized. Since the earthquake data of certain positions have already been recorded in the Taiwan Strait, we then use the above probabilistic model to estimate the disaster level for the unknown areas.

Fig. 5 shows the estimated results with different iterations. We also enlarge the results of the area 22.9N to 24.5N, 118.1E to 120.1E in the right-bottom subfigures. From the simulation results shown in Fig. 5, we can see that the estimated disaster level is propagated from labeled points to their neighborhood points iteratively. This process converged after 200 iterations. Another important observation is that the estimated earthquake-prone area becomes increasingly smooth in a way that the discrete points sampled at different positions are connected together to make a continuous area. This is reasonable as in real-world case the seismic belt is actually smooth and continuous. Hence the results in Fig. 5 provide good seismic belt estimation. In addition, we should also say that this estimation is only utilized for cable routing but not for earthquake prediction.

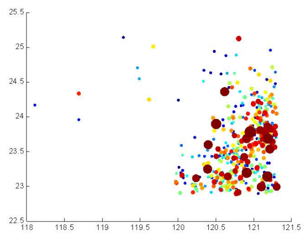
After obtaining the estimated results as in Fig. 6, we can utilize them with three-dimensional topology coordinates to form four-dimensional attributes and to calculate the weights of graph as in Section III-B. Here, we use the Mercator projection method to transform the distance between nearby positions from latitude and longitude representations to meters,



(a)



(b)



(c)

Fig. 4: The topographical map and earthquake data of Taiwan Strait (a) The selected geographic area. The black frame shows the data source of Taiwan Strait. The two red points represent the two endings of the submarine cable; (b) the three-dimensional topographical map. Points with darker color on the map imply deeper areas under the Taiwan Strait Ocean; (c) the earthquake data record in Taiwan, in which each point represents a center of an earthquake. The size of a point is proportional to the corresponding earthquake magnitude and we use different colors to distinguish the overlapped earthquake centers.

which is consistent with the measures of the height for telecommunication cabling. We then search the shortest route along the graph for telecommunication cabling.

The simulation results are in Fig. 6, where the first two columns show the shortest routes across the rough terrain in three-dimensional topology and two-dimensional contour surface, respectively, while the last column shows those across the earthquake area. The shortest routes are all denoted by red line. As proposed in Section III-C, there are two main parameters γ and λ for balancing the tradeoff between the earthquake disaster level and the cabling cost bypassing a mountain or valley. For the convenience of comparison, we first let $\gamma = 0$ and $\lambda = 0$, i.e., the earthquake factor and the landform factor are not taken into account when searching the shortest route. From Fig. 6(a), we can see that the shortest route found by setting $\gamma = 0$ and $\lambda = 0$ will not consider the terrain of the real seafloor and earthquake effect. As shown in the results, the shortest route may cross many rough areas (such as valleys or mountains peaks) and earthquake-stricken area, which is not practical in real-world undersea cabling design. On the other hand, if we let $\lambda = 0$ and set γ to be a large value (such as $\gamma = 100$), we observe that the shortest route found in Fig. 6(b) can well bypass the mountains or valley in the seafloor, which verifies that the proposed algorithm can search the shortest route on a flat landform.

We next show that the proposed method can also find the shortest route that bypasses the earthquake-prone area. The results are shown in Fig. 6(c) and Fig. 6(d). We first let $\gamma = 0$ and $\lambda = 100$, which means only the earthquake factor is taken into account for cabling. In such a case, the shortest route found in Fig. 6(c) can well bypass the earthquake-prone area in a way that the route only pass through the positions with the low levels of earthquake disaster while avoid those with high levels. In addition, if we further set $\gamma = 100$ and $\lambda = 100$, the shortest route found in Fig. 6(d) can be better, as both the tough area and the earthquake-prone area can be avoided. Another observation is that the routes obtained in Fig. 6(d) are significantly longer than that in Fig. 6(a), which may increase the cost for undersea fiber cabling. But in practice, we can adjust the parameters in the proposed algorithm to achieve a balancing result, such that the cabling distance and earthquake risk factor can be optimized altogether.

V. Conclusion

We have proposed a methodology for searching the optimal undersea cable routing by considering cost minimization and earthquake survivability at the same time. In order to evaluate the earthquake disaster level to maintain high cable survivability, we have used a semi-supervised kernel density estimation model for the likelihood of earthquake disaster. Based on this model, we then presented a new approach on combining an undersea three-dimensional topography with hazard information to construct the graph. Optimal solutions in terms of a tradeoff between cable length and cable survivability can then be found by finding the shortest distance on such a graph. Extensive simulations based on a synthetic dataset and a real-world dataset, formed by the topography and earthquake information of the Taiwan Strait, verify the effectiveness of the proposed method. The results have demonstrated that the proposed methodology can provide an effective way in determining undersea cable routing from a totally new perspective. The proposed computational method is effective and not computationally demanding. At last, we believe the proposed methodology can also be applied to other similar kinds of problems such as pipeline, railway line and highway routing.

References

[1] N. Ganganath, C. T. Cheng, and C. K. Tse, “A constraint-aware heuristic path planner for finding energy-efficient paths on uneven terrains,” *IEEE Trans. Ind. Informat.*, vol. 11, no. 3, pp. 601–611, 2015.

[2] V. Roberge, M. Tarbouchi, and G. Labonte, “Comparison of parallel genetic algorithm and particle swarm optimization for real-time uav path planning,” *IEEE Trans. Ind. Informat.*, vol. 9, no. 1, pp. 132–141, 2013.

[3] A. I. . Sabbah, A. El-Mougy, and M. Ibnkahla, “A survey of networking challenges and routing protocols in smart grids,” *IEEE Trans. Ind. Informat.*, vol. 10, no. 1, pp. 210–221, 2014.

[4] M. Sridharan, M. V. Salapaka, and A. K. Somani, “A practical approach to operating survivable wdm networks,” *IEEE J. Sel. Areas Commun*, vol. 20, no. 1, pp. 34–46, 2002.

[5] D. Zhou and S. Subramaniam, “Survivability in Optical Networks,” *IEEE Netw.*, vol. 14, no. 6, pp. 16–23, 2000.

[6] O. Gerstel and R. Ramaswami, “Optical layer survivability: a services perspective,” *IEEE Commun. Mag*, vol. 38, no. 3, pp. 104–113, 2000.

[7] L. M. Capisani and A. Ferrara, "Trajectory planning and second-order sliding mode motion/interaction control for robot manipulators in unknown environments," *IEEE Trans. Ind. Electron.*, vol. 59, no. 8, pp. 3189–3198, 2012.

[8] H. Kim and B. K. Kim, "Online minimum-energy trajectory planning and control on a straight-line path for three-wheeled omnidirectional mobile robots," *IEEE Trans. Ind. Electron.*, vol. 61, no. 9, pp. 4771–4779, 2014.

[9] J. Chan. (2007) Taiwan earthquake triggers a digital tsunami in asia. [Online]. Available: https://www.wsws.org/en/articles/2007/01/taiw-j09.html

[10] C. Barras. (2009) Why the mediterranean is the achilles heel of the web. [Online]. Available: https://www.newscientist.com/article/dn16394-why-the-mediterranean-is-the-achilles-heel-of-the-web/

[11] USGS. (2011) Magnitude 9.0 near the east coast of honshu, japan. [Online]. Available: http://earthquake.usgs.gov/earthquakes/eqinthenews/2011/usc0001xgp.php

[12] K. S. Ho and K. W. Cheung, "Generalized survivable network," *IEEE/ACM Trans. Netw*, vol. 15, no. 4, pp. 750–760, 2007.

[13] A. Narula-Tam, E. Modiano, and A. Brzezinski, "Physical topology design for survivable routing of logical rings in wdm-based networks," *IEEE J. Sel. Areas Commun*, vol. 22, no. 8, pp. 1525–1538, 2004.

[14] I. Ouveysi, A. Wirth, A. Yeh, and O. Oğuz, "Large scale linear programs and heuristics for the design of survivable telecommunication networks," *Annals of Operations Research*, vol. 124, no. 1-4, pp. 285–293, 2003.

[15] I. Ouveysi and A. Wirth, "On design of a survivable network architecture for dynamic routing: Optimal solution strategy and an efficient heuristic," *European Journal of Operational Research*, vol. 117, no. 1, pp. 30–44, 1999.

[16] X. Wang, X. Jiang, and A. Pattavina, "Assessing network vulnerability under probabilistic region failure model," in *Proc. IEEE HPSR*, 2011, pp. 164–170.

[17] P. K. Agarwal, A. Efrat, S. K. Ganjugunte, D. Hay, S. Sankararaman, and G. Zussman, "The resilience of wdm networks to probabilistic geographical failures," *IEEE/ACM Trans. Netw*, vol. 21, no. 5, pp. 1525–1538, 2013.

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

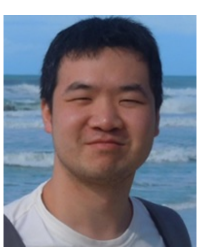
[19] (2016) President's 2016 budget proposes $1.2 billion for the USGS. Directions Magazine. [Online]. Available: http://www.directionsmag.com/pressreleases/presidents-2016-budget-proposes-1.2-billion-for-the-usgs/433229

[20] C. Cao, "Cost effective and survivable cabling design under major disasters," Ph.D. dissertation, Department of Electronic Engineering, City University of Hong Kong, 2015.

[21] C. Cao, Z. Wang, M. Zukerman, J. H. Manton, A. Bensoussan, and Y. Wang, "Optimal cable laying across an earthquake fault line considering elliptical failures," *IEEE Trans. Rel.*, to appear.

[22] M. Wang, X.-S. Hua, T. Mei, R. Hong, G. Qi, Y. Song, and L.-R. Dai, "Semi-supervised kernel density estimation for video annotation," *Computer Vision and Image Understanding*, vol. 113, no. 3, pp. 384–396, 2009.

[23] W. Van, J. Cees, E. Castellanos, and L. K. Sekhar, "Spatial data for landslide susceptibility, hazard, and vulnerability assessment: an overview," *Engineering geology*, vol. 102, no. 3, pp. 112–131, 2008.

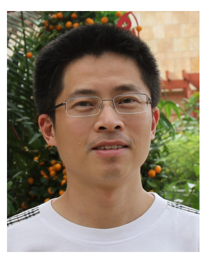


Mingbo Zhao (S'11–M'13) received the BSc and master's degree from the Department of Electronic Engineering, Shanxi University, Shanxi, China, in 2005 and 2008, respectively, and the Phd degree in computer engineering from the Department of Electronic Engineering, City University of Hong Kong, in 2013. He is currently a senior research assistant in the Department of Electronic Engineering, City University of Hong Kong. His interests include machine learning, data mining and pattern recognition.

Tommy W. S. Chow received the B.Sc. (1st Hons) degree and the Ph.D. degree from the Department of Electrical and Electronic Engineering, University of Sunderland, U.K. He is currently a Professor in the Department of Electronic Engineering at the City University of Hong Kong. His main research areas include neural networks, machine learning, pattern recognition, and fault diagnosis. He received the Best Paper Award in 2002 IEEE Industrial Electronics Society Annual meeting in Seville, Spain. He is an author and co-author of over 170 technical Journal articles related to his research, 5 book chapters, and 1 book. He now serves the IEEE Transactions on Industrial Informatics and Neural Processing letters as Associate editor.

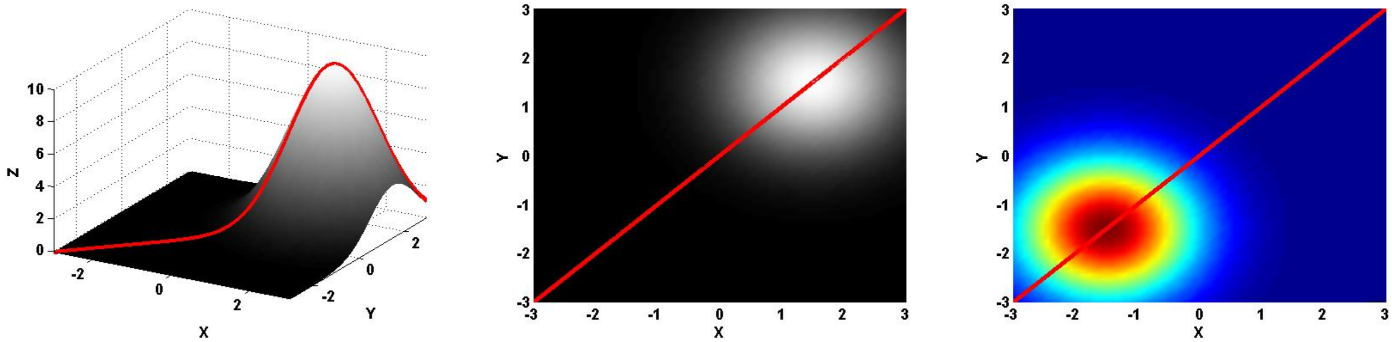
Peng Tang received the Ph.D. degree at the Department of Electronic Engineering, City University of Hong Kong, Hong Kong, in 2014. He received his B.Eng. degree from the Department of Information Science and Technology, Shandong University, China, in 2007, and master degree from the Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing, China, in 2010, respectively. His current interests include document retrieval, computational intelligence, pattern recognition, and their applications.

Zengfu Wang received the B.Sc. degree in applied mathematics, the M.Sc. degree in control theory and control engineering, and the Ph.D. degree in control science and engineering from Northwestern Polytechnical University, Xi'an, China, in 2005, 2008, and 2013, respectively. From 2014 to 2015, he was a Postdoctoral Fellow with City University of Hong Kong. Currently he is a Lecturer with Northwestern Polytechnical University. His research interests include path planning, discrete optimization and information fusion.

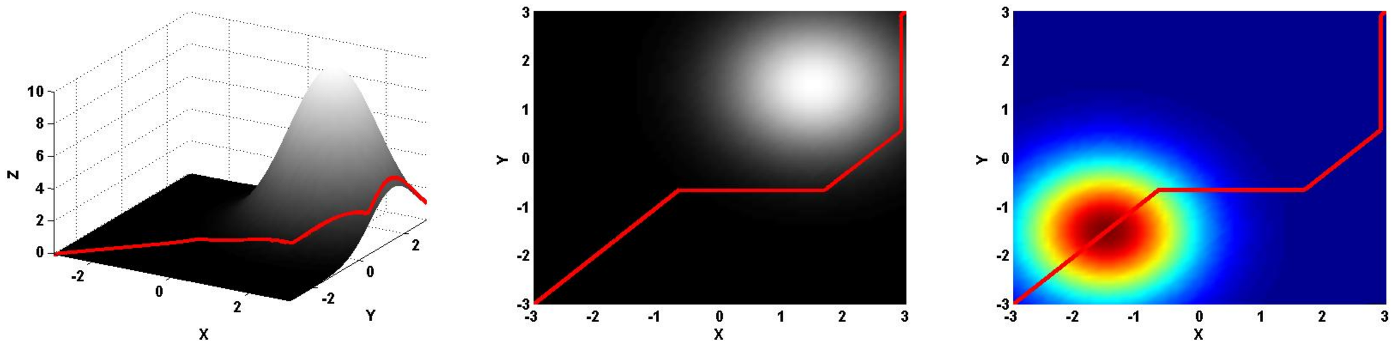


Jun Guo (S'01–M'06) received the B.E. degree in automatic control engineering from Shanghai University of Science and Technology, Shanghai, China, in 1992, and the M.E. degree in telecommunications engineering and the Ph.D. degree in electrical and electronic engineering from the University of Melbourne, Melbourne, Australia, in 2001 and 2006, respectively. He was with the School of Computer Science and Engineering, The University of New South Wales, Kensington, Australia, from 2006 to 2011. Since 2012, he has been with the Department of Electronic Engineering, City University of Hong Kong, where he is currently a Senior Research Fellow. His research interests include green communications and networking, teletraffic theory and its applications in service sectors, and survivable network topology design.

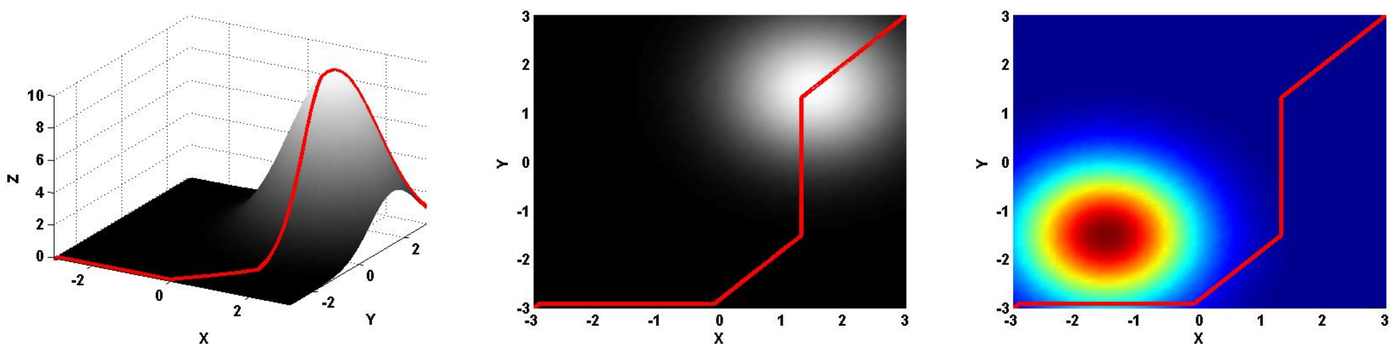
Moshe Zukerman (M'87–SM'91–F'07) received his B.Sc. and M.Sc. degrees from the Technion, Israel, and his Ph.D. degree in engineering from University of California, Los Angeles, in 1985. He was an independent Consultant with the IRI Corporation and a Postdoctoral Fellow with the University of California, Los Angeles, in 1985–1986. During 1986–1997, he was with Telstra Research Laboratories. He also taught and supervised graduate students at Monash University in 1990–2001. During 1997–2008, he was with The University of Melbourne, Victoria, Australia. In 2008 he joined City University of Hong Kong as a Chair Professor of Information Engineering, and a team leader. He has over 300 publications in scientific journals and conference proceedings. He has served on various editorial boards, and contributed to numerous conference technical committees.



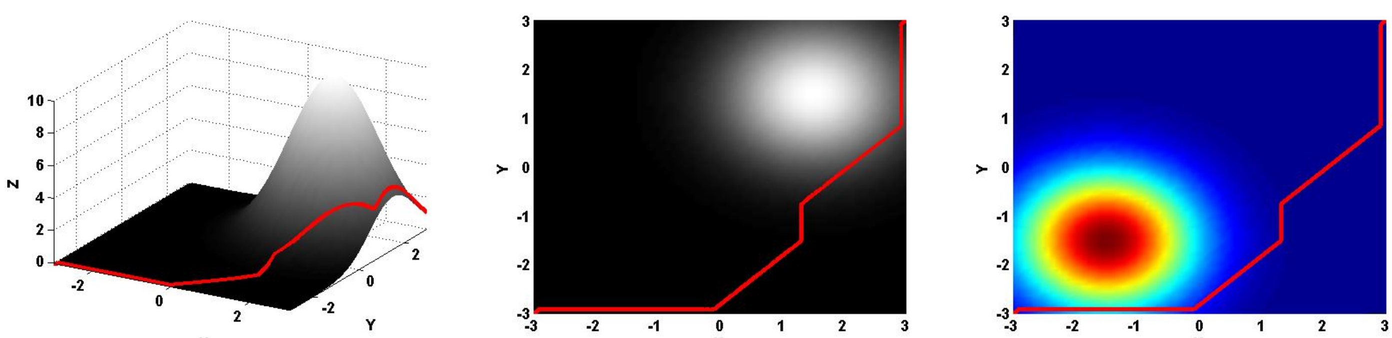
(a) The shortest route learned without considering the factors of terrain and earthquake ($\lambda = 0$ and $\gamma = 0$)



(b) The shortest route learned by only considering the factor of terrain ($\lambda = 0$ and $\gamma = 100$)

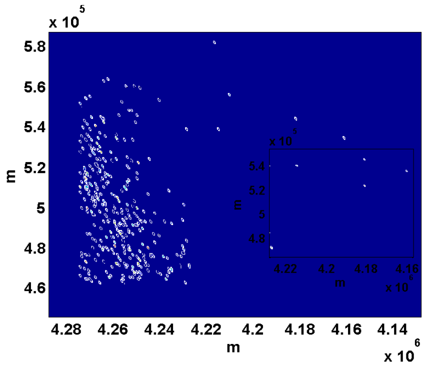


(c) The shortest route learned by only considering the factor of earthquake ($\lambda = 100$ and $\gamma = 0$)

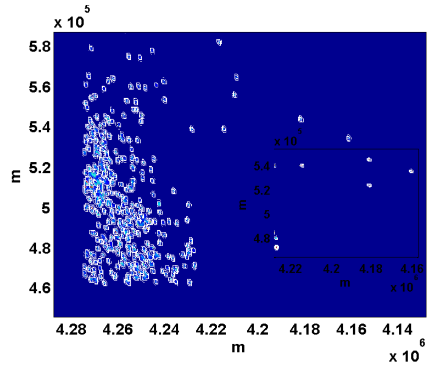


(d) The shortest route learned by considering the factors of terrain and earthquake ($\lambda = 100$ and $\gamma = 100$)

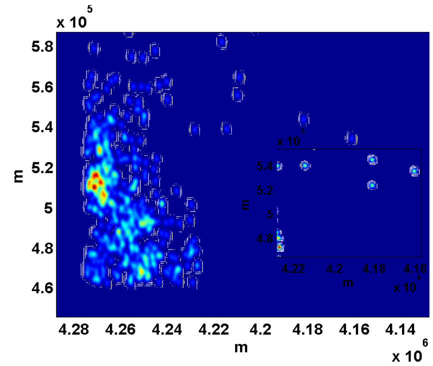
Fig. 3: The shortest route learned on a synthetic dataset: the data is sampled from the region of $\{(X, Y)\,|X \in [-3, 3], Y \in [-3, 3]\}$, where a mountain is generalized around the point $[2.5, 2.5]$ and a earthquake is generalized at $[-2.5, -2.5]$. The shortest route learned in all cases is denoted by the red line. The first two columns show the shortest routes across the rough terrain in three-dimensional topology and two-dimensional contour surface, respectively, while the last column shows those across the earthquake area.



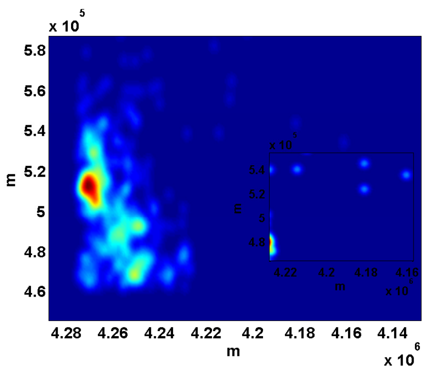
(a) 0 iterations



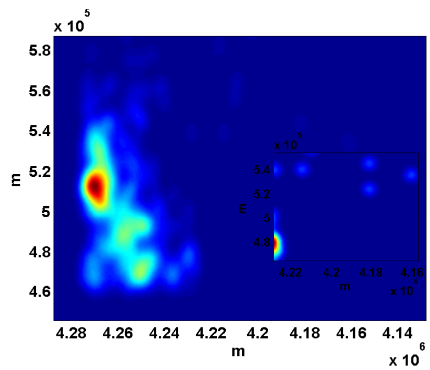
(b) 5 iterations



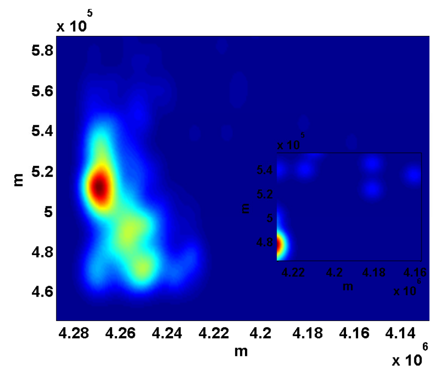
(c) 10 iterations



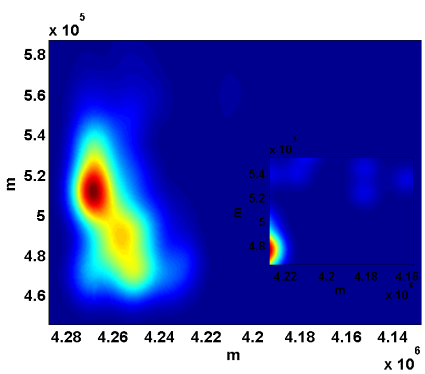
(d) 20 iterations



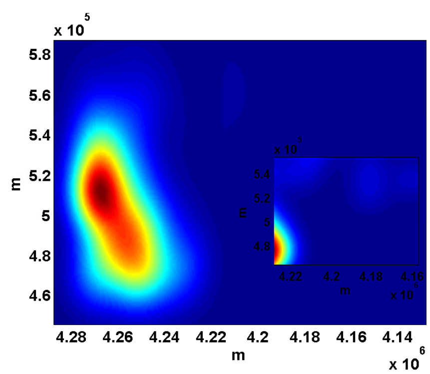
(e) 40 iterations



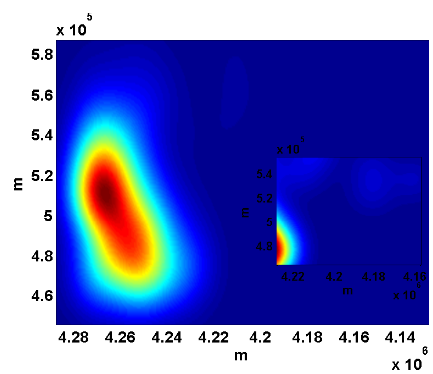
(f) 80 iterations



(g) 120 iterations

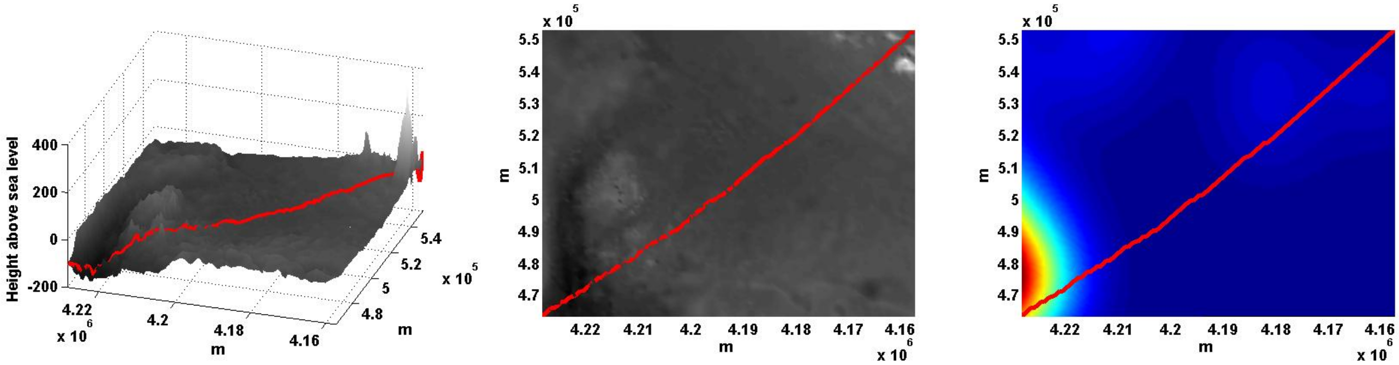


(h) 160 iterations

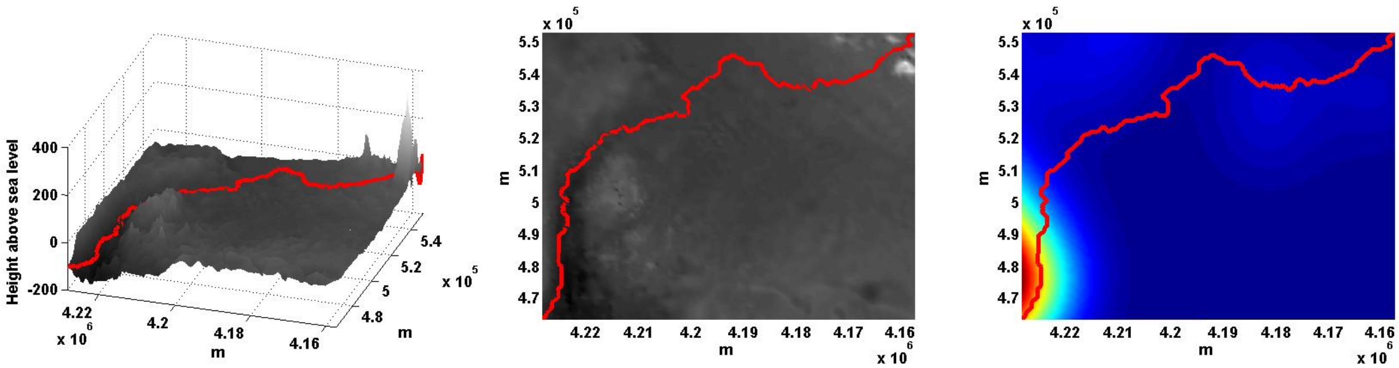


(i) 200 iterations

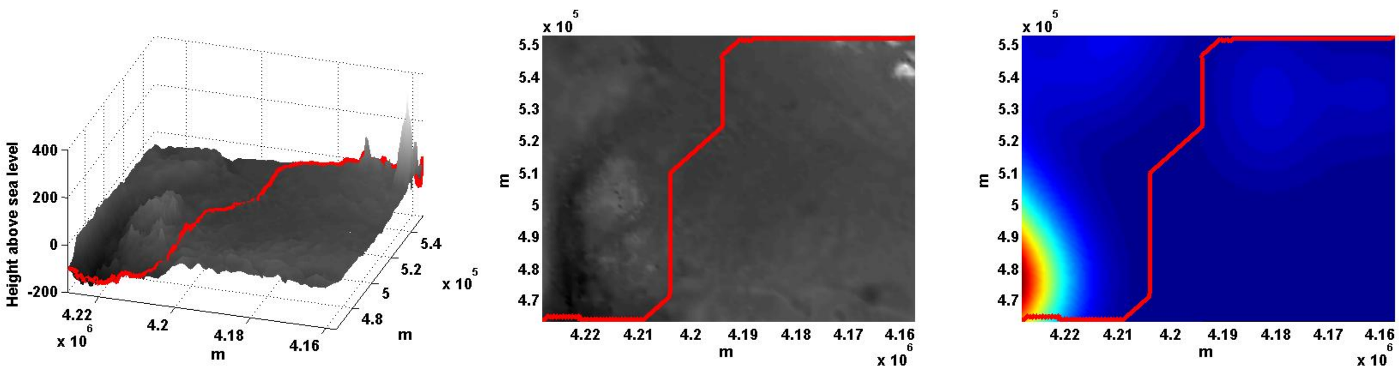
Fig. 5: The estimated likelihood of earthquake with different iterations: Taiwan Strait earthquake data. The results of the area 22.9N to 24.5N, 118.1E to 120.1E are also enlarged in the right-bottom subfigures.



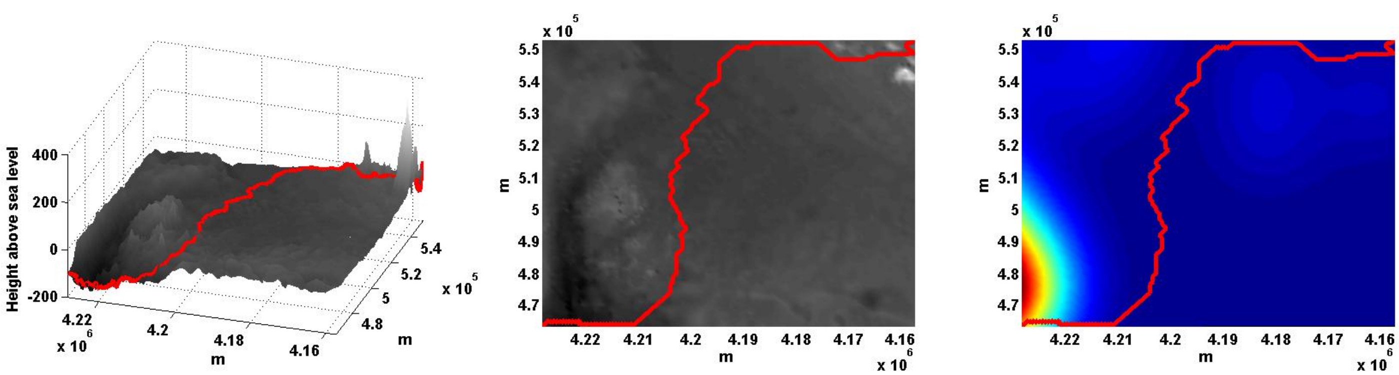
(a) The shortest route learned without considering the factors of terrain and earthquake ($\lambda = 0$ and $\gamma = 0$)



(b) The shortest route learned by only considering the factor of terrain ($\lambda = 0$ and $\gamma = 100$)



(c) The shortest route learned by only considering the factor of earthquake ($\lambda = 100$ and $\gamma = 0$)



(d) The shortest route learned by considering the factors of terrain and earthquake ($\lambda = 100$ and $\gamma = 100$)

Fig. 6: The shortest route learned on real-world Taiwan Strait topology and earthquake data. The shortest route learned in all cases is denoted by red line. The first two columns show the shortest routes across the rough terrain in three-dimensional topology and two-dimensional contour surface, respectively. The last column shows those across the earthquake-prone area.