ARTICLE IN PRESS

Reliability Engineering and System Safety xxx (xxxx) xxx

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

Analyzing vulnerability of optical fiber network considering recoverability

Ke Wang [a], Jinfeng Liu [a], Lai Tian [a], Xianfeng Tan [a], Guansheng Peng [b], Tianwen Qin [a], Jun Wu [c,*]

[a] Communication Non-commissioned Officers Academy, Army Engineering University, Chongqing 400000, PR China
[b] College of System Engineering, National University of Defense Technology, Changsha, 410073, PR China
[c] International Academic Center of Complex Systems, Beijing Normal University at Zhuhai, 519087, PR China

ARTICLE INFO

Keywords:
Optical fiber network
Large-scale damage
Damage measurement
Recoverability
Vulnerability

ABSTRACT

With the development of optical transmission technology, optical fiber networks have become critical infrastructures in supporting information transmission on the Internet. However, the fiber cable is very vulnerable to large-scale damage such as earthquakes or pulse bombs. What is more serious is that it will take a long time to locate and repair the damages on fiber links. The long-term repair process will cause continuous network performance degradation and severe economic loss. The fact is that these dangerous areas may be ignored by traditional vulnerability analysis models. To solve this problem, this paper proposes a method to analyze the vulnerability of fiber networks based on network recoverability. We first improve the traditional fiber network simulation methods and damage simulation methods to provide a compatible foundation for the network recoverability simulation. Then, we present the network vulnerability analysis model: the Damage Measurement and Location Model (DMLM). The model employs the heuristic traversal algorithm based on random points to locate the candidate attack positions. We design three vulnerability metrics: two metrics are related to network recoverability and one metric is used for comparison. We also build their corresponding theoretical frameworks to determine the appropriate model parameters to satisfy the specified estimation error requirements. Numerical results prove the proposed model's effectiveness and excellent sensitivity for essential parameters. The visual results of the vulnerable zones prove the necessity of considering network recoverability in vulnerability analysis of optical fiber networks.

1. Introduction

With the development of computer technology and communication technology, the way of information transmission has changed dramatically. Optical cable has quickly replaced the traditional copper cable, becoming the core facility of the communication backbone network. Compared with traditional fixed communication materials, optical fiber has the following advantages: reliable transmission technology of wide frequency band, large capacity, and low loss of transmission, as well as the commercial value of low cost, small weight and long service life. In China, the total length of cable lines has reached 45.46 million kilometers as of June 2019. Submarine optical cables around the world have accumulated more than 1.4 million kilometers and carry 98% of the bandwidth of international communications. People's daily life, operation of enterprises, and nation-building are all increasingly inseparable from the basic information services provided by optical fiber networks. However, the optical cable has obvious defects in damage resistance: low mechanical strength, high bending angle requirement, and high continuous power supply requirements [1,2]. When confronting natural disasters (such as earthquakes [1], tsunamis [3], typhoons, et cetera.) or intentional attacks (such as bomb attacks and electromagnetic pulse bomb attacks [4]), optical cables are prone to be damaged, distorted or power outage. Earthquake located in Taiwan Strait on December 26, 2006, damaged seven submarine cables of China–US Cable Network, blocked the normal communication service between China and the US for two weeks [5]. The Great East Japan Earthquake on March 11, 2011, and the tsunami triggered by earthquake-damaged approximately 1.5 million circuits for fixed lines and 4900 mobile base stations [6].

It should be vigilant that a disaster or attack will not only immediately cut off the information transmission in the optical fiber network, but also bring more serious problems: it will take a lot of time, manpower and material resources to locate and repair damaged positions of the fault fibers. In a complex post-disaster environment or undersea environment, the fault location and repair speed will be very low, which can easily cause persistent communication barriers and economic losses. The Taiwan Earthquake took repair ships 48 h to locate the fault in the submarine cables, and one month to complete

* Corresponding author.
E-mail address: junwu@bnu.edu.cn (J. Wu).

https://doi.org/10.1016/j.ress.2021.108308
Received 23 July 2021; Received in revised form 29 October 2021; Accepted 26 December 2021
Available online 10 January 2022

the maintenance. In the Great East Japan Earthquake, 90% of communication capacity was restored after 75 days through an all-out effort by over 11,000 people.

Although researchers have shown great interest in analyzing the vulnerability of communication networks under geographical failures during the last two decades [7–9], most studies ignore the impact of network repair difficulty when analyzing network vulnerabilities. If the vulnerability analysis is only performed according to the network connectivity or reliability, it is easy to miss some special dangerous areas. Therefore, this paper is devoted to analyzing the optical fiber network vulnerability while considering network recoverability, thus providing a new perspective for network protection and maintenance.

The biggest problem to be solved in this research direction is the compatibility of the fiber recoverability simulation model and the fiber network vulnerability analysis model. The fundamental reason for the incompatibility lies in the lack of a series of basic simulation models that can uniformly measure and accurately describe the specific damage location, damage degree, repair difficulty, and repair progress of the fiber link. The specific reasons are as follows.

1. Fiber network vulnerability analysis models. There are two main problems: the first is that the simulation accuracy of the existing fiber link model does not meet the calculation requirements; the second is that the traditional damage simulation model cannot directly and accurately describe the network's actual damage degree.

Many fiber network models treat the optic cable as a simple whole for simulation, and this type of simulation model only includes the coordinates of both ends of the optic cable. Since the optical cable is generally tens of kilometers long, when damage occurs, the damage degree is different for different damage locations. A simple model cannot meet the simulation requirements. Therefore, we designed the Link Simulation Model (LSM). The model solves the simulation problem of each location of the fiber link by dividing the link evenly. The model provides a basic platform for fine simulation of fiber cable damage and fine simulation of fiber cable repair.

To solve the damage-efficiency problem, we design the Damage-efficiency Circle Model (DECM). The damage-efficiency refers to the comprehensive judgment index of the damage degree of the target under the natural disaster or attack [10]. Target damage probability and target performance loss expected value are both commonly used indicators to describe target damage-efficiency. Different from the simulation that uses probability to describe the damage-efficiency of fiber cables [5,11,12], the damage-efficiency in this model describes the actual damage degree to the network equipment. That is, the expected value of the loss of the fiber network's transmission capacity after damaged. This method can ensure that devices with different transmission capabilities have the same capability loss when their distance from the damage center is the same. The damage probability model must assume that the transmission capacity of each device is the same to meet the above requirements, which does not match the actual situation. In the damage simulation model, we employ circles [11,12] to simulate the shape of geographical damage. The damage-efficiency of nodes and virtual points is a function of their distance to the circle center.

2. Fiber network recoverability models. The network recoverability refers to the ability of the fiber network to restore its data transmission capacity to or close to the level before damage through repair after damaged. The main problem is that the simulation accuracy of the fiber network repair process is insufficient. Most of the existing literature focuses on the optimal design of the fiber network repair strategy, so the accurate simulation of the damage is not essential. The simplification of the damage simulation results in the simplification of network repair process. Based on the Link Simulation Model and the Damage-efficiency Circle Model, we designed a repair simulation model with higher accuracy and closer to reality. We also designed a simple, applicable, and relatively close to the actual repair strategy according to the characteristics of the fiber network setting: the Saturated Rescue Strategy (SRS).

Saturated rescue means that when a country or region suffers a major natural disaster or deliberately attack, the government quickly and macro-mobilizes the local and surrounding resources to implement emergency rescue measures including medical treatment, food, electricity, and communications in the disaster-affected area [13]. This is a common disaster emergency rescue strategy in many countries, and it is also an emergency rescue strategy advocated by the United Nations. The Saturated Rescue Strategy for fiber networks refers to that after a single damage to the fiber network, there must be a sufficient number and well-equipped maintenance teams to perform the repair, and the repair resources are sufficient during the repair process to ensure that the average repair speed of the rescue team reaches the required level and remains unchanged. The technical concepts mentioned in this strategy are mainly derived from the telecom operators' actual emergency repair process after the cable damaged [2,6,14] and the optical cable repair manual used by the author's college teaching. The vulnerability analysis model using this strategy no longer requires additional input of repair strategy data, which significantly simplifies the complexity of the model.

After perfecting the compatibility of the basic models, we have designed the fiber network vulnerability analysis model, the Damage Measurement and Location Model (DMLM). The traversal algorithm based on random points from Wang [5] is adopted to search network vulnerability locations. Compared with the computational geometry method, this method has higher calculation accuracy and can analyze and prove the error range of the result. However, the calculation complexity of the original algorithm is too high. We devised a heuristic method which combined the matrix and computational geometry methods to optimize the original traversal algorithm to make the calculation time acceptable. Subsequently, we designed two network performance metrics in combination with network vulnerability and recoverability. They are the Maximum Repair Time (MRT) and the Continuous Performance Degradation (CPD). To show the changes in the vulnerability analysis results after considering network recoverability, we imitate the vulnerability metrics used in both Wang's and Agarwal's models: the total remaining link capacity, and design Remain Transmitting Capacity Ratio (RTCR). Based on the three metrics, the calculation formulas of the model parameters under specified estimation error requirements are proposed with analytic proofs.

The main contributions of this paper are as follows:

(1) A vulnerability analysis model of fiber network based on network recoverability is developed.

(2) Multiple vulnerability analysis metrics considering the network repair process are designed.

(3) The formulas of important model parameters under the given error requirements are proved analytically.

(4) Two case studies are used to examine the validity, sensitivity, and practicability of the proposed model.

(5) Real case study demonstrates that considering network recoverability will significantly affect the identification of fiber network vulnerable areas.

The rest of this paper is organized as follows: Section 2 summarizes and analyzes the related researches. Section 3 describes the simulation models of fiber network, damage and recoverability. Section 4 presents the vulnerability analysis model and the three vulnerability metrics. Section 5 presents the numerical results. The last section gives a conclusion and briefly discusses future challenging topics.

2. Literature review

This section summarizes and analyzes the research literature in three related fields: fiber network recoverability, network resilience, and network vulnerability analysis.

2.1. Related research on fiber network recoverability

Many researches on the recoverability of fiber networks focus on the discussion of repair strategies, such as progressive network recovery strategies under resource-constrained conditions [15–17]. These studies first assume that network repair resources are relatively scarce, and only part of the repair can be completed within the required time. Therefore, heuristic algorithms need to be designed according to different priority requirements (critical link first, key node first, or key business flow first). Some studies pay more attention to how to effectively formulate network repair strategies under uncertain damage scenarios [18]. Some researches focus on the design of repair strategies [19] and innovations in repair methods [20] for the failure of the access part of the fiber network. Some studies focus on how the restoration entities (repair workers, repair vehicles) choose the optimal path of movement and how to optimize the allocation of repair resources [21,22]. Unfortunately, these studies pay more attention to topological analysis. Even considering the spatial properties of the network, the optical cable is modeled as a simple coordinate unit. Due to the high computational complexity of the strategy optimization model, most studies are difficult to distinguish the specific damage location and damage degree of the optical fiber link. The difficulty of algorithm implementation can only be reduced by reducing the simulation accuracy of the repair process.

2.2. Related research on network resilience

It is worth noting that the research on network resilience will also simulate the network damage condition and network repair process. In particular, when we conduct network vulnerability analysis based on recoverability, the model and metric design are similar to network resilience modeling [23,24]. However, the definition of system resilience includes not only the ability of the system to recover to its original performance after being damaged, but also the ability of the system to resist disturbance before and during the damage [25,26]. System resilience is consistent with system vulnerability in positioning, which is a metric to comprehensively measure system performance from the macro perspective [27]. System recoverability is one of the many factors that affect system resilience or system vulnerability. The current research on system resilience is still in the exploration stage [28]. Most of the research focuses on the resilience modeling of infrastructure such as transportation networks, power networks, and water conservancy networks [29,30]. In the direction of optical fiber transmission system, there are fewer clear metric modeling methods. Therefore, this paper does not use system resilience as an indicator to describe the network repair situation.

2.3. Related research on network vulnerability analysis

Most of the researches are to analyze the network vulnerability by attacking the network and observing the degradation of network performance, the so-called Network Inhibition Problem [31,32]. The modeling of network vulnerability analysis model needs to include three parts: network simulation model, damage simulation model, and optimal damage area search algorithm. We introduce the relevant literature one by one according to these three parts.

For fiber network modeling, most researches simulate the fiber link as a simple unit [33–35]. In other words, the link will be regarded as a whole for coordinate marking, and it will also be regarded as a whole for marking the damage degree. If the fiber link is regarded as a simple unit during the repair process, the calculated network vulnerability and the found vulnerability areas will be seriously distorted in a mass destruction scenario.

For damage simulation modeling, these researches could be divided into two types in terms of simulation of region failures: deterministic failure models [36–38] and probabilistic failure models [5,11,12]. The deterministic failure model means that the network equipment in the failure region is immediately and completely destroyed. The probabilistic failure model means that the network equipment in the failure region is destroyed with a certain probability, and the destroyed probability is a function of the equipment's distance to the failure center. In Agarwal's model [11], they designed a probabilistic hippodrome to pre-process the network, then found the attack position based on the superposition result of the probabilistic hippodromes. The probabilistic hippodrome only provides a calculation basis for searching reasonable attack locations. Fiber links are still regarded as a whole in the final network vulnerability analysis. In Wang's model [5], they designed a segmented probability damage model: each probability segment is equivalent to a concentric circle, and the damage probability of the fiber link in the same concentric circle is the same. As long as the segmentation range is small enough, a satisfying simulation for link damage can be achieved. However, they only gives $2 \sim 3$ probability segments in the actual network analysis, indicating that this method is difficult to meet the simulation accuracy requirements due to its high complexity. No matter how the probabilistic damage model is designed, there will be an inevitable defect: it essentially describes whether the network equipment is destroyed, not the loss of equipment transmission capacity. This is inconsistent with the actual network damage: even if the equipment survives with a high probability, it will likely lose some functionality. The probabilistic damage model cannot describe this scenario due to its definition.

For damage area researching, because the network has spatial attributes, computational geometric methods and traversal algorithms are the most common methods to search for the best attack location. Some researches set the specific position of the damage line segment and the damage circle through the angle between the links and the position of the node [33,39]. This method provides the proof method to determine the best damage location, but the computational complexity is relatively high. Some researches set the damage circle position by determining the minimum enclosing circle of the specified node [36,40]. The algorithm complexity of this method is low, but the accuracy of the calculation result is not guaranteed. Some researches set the damage circle position through the seismic hazard maps [12]. This method is limited to describing earthquake disasters. Some researches determine the location of the damage circle by traversing evenly distributed grids on the two-dimensional plane where the network located [5]. This method can analytically calculate the upper limit of the grid diameter under a given error requirement, ensuring the accuracy of the calculation result, but the calculation complexity is high.

3. Simulation modeling for fiber network, damage and recoverability

Because the traditional network simulation model, damage simulation model and recoverability simulation model have compatibility problems, we redesign the optical fiber network modeling method and damage simulation method and propose the Link Simulation Model (LSM) and the Damage-efficiency Circle Model (DECM). New methods can improve the accuracy of model simulation and provide a calculation basis for simulation modeling of recoverability. At the end of this section, we detail the simulation model of network recoverability with the Saturated Rescue Strategy (SRS).

3.1. The link simulation model

In optical fiber networks, communication base stations, fiber switches, and routers are commonly abstracted into nodes, denoted as v_i. Optical cables (including amplifiers) are commonly abstracted into links, denoted as e_{ij}. The topology graph $G = (V, E)$ describes the topology of the network, where V is the set of nodes and E is the set of links. Network's nodes are located in a two-dimensional Euclidean plane, and the coordinates of v_i can be denoted as (x_i, y_i). If the original data of (x_i, y_i) uses latitude and longitude coordinates, to ensure the accuracy of the calculation results, we recommend using the Gauss–Kruger Projection to convert the latitude and longitude coordinates into two-dimensional plane coordinates.

For the simulation of the fiber backbone cable, most researches simplify it to a straight line segment with nodes at both ends [5,11,37]. The main reason is: affected by its own material, the optical cable between adjacent base stations on land is generally laid (erected) in a straight direction. The fiber material has a very low tolerance to longitudinal traction, and dragging must be avoided when laying. If the laying route has curvature, the longitudinal traction will be greatly increased due to friction. At the same time, the laying angle of the optical fiber should not be too large, and multiple small angle bends should be avoided in a short distance, otherwise it will seriously affect the light transmission efficiency. Therefore, we also simulated the backbone optical cable as a straight line segment, and the link e_{ij} can be uniquely located by its end nodes v_i and v_j.

The information transmitting capacity of the optical fiber network is generally measured by its bandwidth. The bandwidth of the fiber network node generally refers to the port bandwidth or backplane bandwidth of the network relay equipment (such as the fiber switch). The backplane bandwidth of the relay equipment used by network operators has a large gap, but the port bandwidth is basically similar, generally tens of gigabit per second. Therefore, we choose the port bandwidth as the node transmitting capacity, and the transmitting capacity of v_i is denoted as f_i. The theoretical bandwidth value of the optical fiber link is tremendous. However, the actual bandwidth is limited by the photoelectric conversion technology of the relay equipment and is generally slightly smaller than the node port bandwidth. The transmitting capacity of the link e_{ij} is defined as f_{ij}, and there is $f_{ij} \approx \mathbf{min}\{f_i, f_j\}$. In order to reduce the complexity of the overall model, at the same time eliminate the influence of different units of bandwidth (gigabit per second or megabits per second), we replace f_i and f_{ij} with the ratio of f_i and f_{ij} with the maximum transmission capacity in the network:

$$F_i = f_i/\mathbf{max}\{f_1, \dots, f_{n_v}\}, F_{ij} = f_{ij}/\mathbf{max}\{f_1, \dots, f_{n_v}\}. \tag{1}$$

Where n_v is the number of nodes. It should be mentioned that if model users need to compare and analyze the vulnerabilities of different networks, the denominator of Formula (1) is supposed to be the maximum value of the transmission capacity of all nodes of all networks.

Optical fiber could be regarded as seamless welding of several shorter fibers with the same length. When the length of the shorter fiber is small enough, each position of the shorter fiber will have a close geographical location, running condition, and damage efficiency. As long as the damage efficiency of one position is calculated, the status of others can be obtained directly. For simplifying the model, the midpoint of the shorter fiber is selected as the representative, which is called Virtual Point, and denoted as p_k^{ij}. Where i, j explain which link e_{ij} the virtual point belongs to and k expresses that this is the kth virtual point of this link. The segmentation distance, i.e., the length of the shorter fiber, is denoted as s. As the value of s decreases, the simulation error of fiber cables will decrease, but the simulation error is bound to exist objectively. We combined the network vulnerability metrics to discuss the value range of s under the given error requirements in Section 4.

The properties of virtual points p_k^{ij} of the link e_{ij} could be described by the matrix as follows:

$$\begin{pmatrix} x_1^{ij} & y_1^{ij} & i & j & \varphi_e & F_1^{ij} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_m^{ij} & y_m^{ij} & i & j & \varphi_e & F_m^{ij} \end{pmatrix}. \tag{2}$$

The matrix has m rows and 6 columns. m is the number of virtual points in link e_{ij}, which is calculated as $m = \mathbf{Int}\left(l_{ij}/s\right)$, where $\mathbf{Int}$ is a floor rounding function, and l_{ij} is the length of the link e_{ij}. x_k^{ij}, y_k^{ij} are the coordinates of the virtual point p_k^{ij}, and i, j are the index values of nodes v_i, v_j in the set V. φ_e expresses the invulnerability of the link's equipment. The invulnerability of the link is comprehensively affected by multiple factors such as its laying method and protective materials. In order to simplify the model, we have not discussed and analyzed in depth, and only designed φ_e to represent the comprehensive performance value of the link invulnerability. The way the same network operator installs the optical cable and the protective materials used are the same, so we also normalize the φ_e of all links, there is $\varphi_e \in [0, 1]$. The smaller the value of φ_e, the worse the ability of the equipment to resist damage.

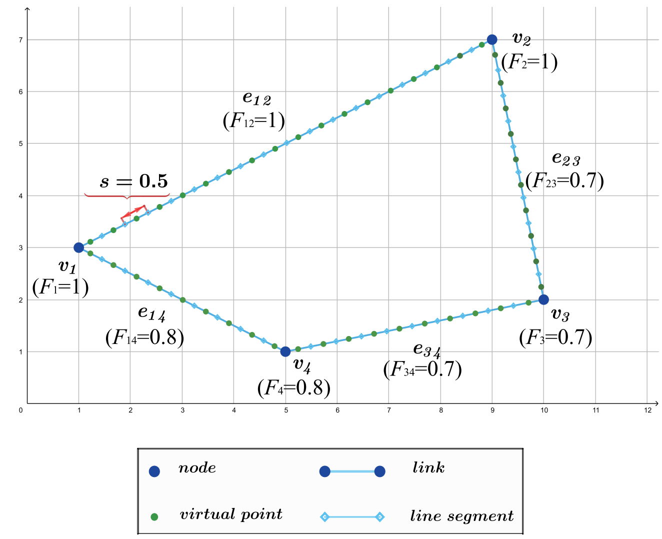

Fig. 1. Results of virtual points calculated by LSM on the example network. Network is embedded on the Euclidean two-dimensional plane. The coordinates of each node are: $(1,3)$, $(9,7)$, $(10,2)$, $(5,1)$. The length of the average line segment s is set as 0.5.

F_k^{ij} indicates the transmitting capacity of the fiber at that location, and its range is $F_k^{ij} \in [0, 1]$. When fiber links suffer damage, F_k^{ij} may decrease. If F_k^{ij} remains unchanged, indicating this position is undamaged and operating normally. If F_k^{ij} is reduced to 0, indicating this position is completely damaged and cannot function. If F_k^{ij} decreases but does not reach 0, indicating this position is not completely damaged but losing some functions, such as the bandwidth loss caused by a few fibers broken in the cable.

Since the virtual point p_k^{ij} is the midpoint of each small segment, the coordinate value of p_k^{ij} can be calculated according to the above matrix as follows:

$$x_k^{ij} = x_i + \left(x_j - x_i\right)\frac{2k-1}{2m}, y_k^{ij} = y_i + \left(y_j - y_i\right)\frac{2k-1}{2m}, \tag{3}$$

where (x_i, y_i) is the coordinate of one end of e_{ij}, (x_j, y_j) is the coordinate of the other end, and there is $k = 1, 2, \dots, m$.

For uniforming the calculation format, the 1×6 metric is presented to describe the properties of the node v_i:

$$\begin{pmatrix} x_i & y_i & i & i & \varphi_v & F_i \end{pmatrix}, \tag{4}$$

where i is the index value of the node v_i, and F_i is the node's transmitting capacity, its range is $F_i \in [0, 1]$. φ_v expresses the invulnerability of the node's equipment. It is generally believed that nodes' invulnerability is stronger than optical fiber links in the fiber network, i.e., $\varphi_v > \varphi_e$.

Fig. 1 illustrates the processing results of the LSM on the example network. We simulated each link in the example network with LSM, and s is the length of the line segment.

3.2. The damage-efficiency circle model

The damage-efficiency refers to the comprehensive judgment index of the damage degree of the target under the natural disaster or attack.

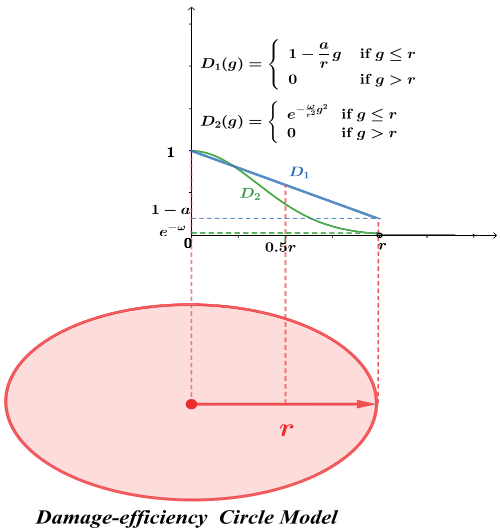

Fig. 2. The Damage-efficiency Circle Model. The bottom half of the graph is the damage circle, and the top half is the variation of the two damage efficiency functions with the damage radius.

Most studies generally judge the damage effectiveness of a single target according to the probability of damage. Obviously, the optical cable is not suitable as a single target to measure the damage degree. Firstly, an optical cable generally contains 8 to 64 fibers, and each fiber works independently. The damage degree of the fiber cable is closely related to the number of damaged fibers. The probability damage model is difficult to describe the loss of cable's transmission capacity caused by the damaged fibers. Secondly, the length of an cable is generally more than 10 kilometers, and the length of the backbone optical cable can even reach 200 kilometers. Such a large geographical span cannot be used as a single target to be calculated in the damage probability model. Therefore, we employ another definition of damage-efficiency: the expected value of the loss of the target's operational capability after being damaged. This definition is mostly used for target individuals or collections with complex internal structures and multiple operating mechanisms. Therefore, the damage-efficiency in our model refers to the expected value of the transmission capacity loss of the optical fiber network after damaged.

The damage circle was always employed to simulate regional failures caused by earthquakes, hurricanes, weapons of mass destruction, or electromagnetic pulse bombs [4,7,41,42]. These disasters or attacks' destructive generally reach its peak at the epicenter and decrease as the damage radius increases. Therefore, the damage efficiency should be a decreasing function, and its independent variable is the damage position's distance to the damage circle center. The simulation calculation of damage-efficiency is very complicated. We have followed the literature [4,11,43] to use the linear function and the probability density function of Gaussian distribution function (hereinafter referred to as the Gaussian distribution function) to simulate damage-efficiency. Different from the traditional function value representing the probability of damage, the function value calculated by our model represents the expected value of the network transmission capacity reduction caused by the attack. Since the network transmission capacity values have been normalized in LSM, the writing formats of the model functions are the same as the traditional functions.

The linear function is described as follows:

$$D_1(g) = \begin{cases} 1 - \frac{a}{r} \cdot g & g \le r \\ 0 & g > r \end{cases} \tag{5}$$

where r is the radius of damage circle, and a is the parameter to adjust the damage degree at the circle edge. Its value range is $a \in [0, 1]$. The larger the value of a, the more clearly the damage attenuation decreases with the damage radius increasing. g is the distance from the node to the damage circle center.

The probability density function of Gaussian distribution is described as follows:

$$D_2(g) = \begin{cases} e^{-\frac{\omega}{r^2} \cdot g^2} & g \le r \\ 0 & g > r \end{cases} \tag{6}$$

where r and g are denoted the same with Formula (5). ω is also the parameter to adjust the damage degree at circle edge, and there is $\omega > 0$. Fig. 2 shows how $D(g)$ changes with the position of the damage circle. When the network equipment is damaged, the change in its transmitting capacity is related to its suffered damage efficiency and its equipment's invulnerability. We assume that the equipment's invulnerability could affect the actual damage efficiency of the equipment, then the calculation formula of the damaged equipment's transmission capacity can be abstracted as follows.

$$F^* = F - (1 - \varphi_v) \cdot D(g), F_p^* = F_p - (1 - \varphi_e) \cdot D(g_p), \tag{7}$$

where F^*, F_p^* are the transmitting capacity of nodes and virtual points after failure, respectively. g_p describes the distance from the virtual point to the damage circle center. There is $F^* \in [0, F]$ and $F_p^* \in [0, F_p]$. $F^* < 0$ or $F_p^* < 0$ may occur in actual calculations, which means that the node or link has suffered more than its transmission capacity. However, the negative value of transmitting capacity is meaningless, so the negative value will be adjusted to 0 in the program calculation.

3.3. The recoverability simulation model

Network recoverability is mainly affected by the difficulty of the repairing process after the network is damaged. Network repair is a complex process, including the repair of damaged entities (large switches, program-controlled computers, optical cable bodies, and optical cable connectors), as well as error correction, adjustment, or resetting of optical network communication programs. Due to space limitations, this article only considers the repair of damaged entities in the network, and the network repair process is simplified to the repair process of network nodes and edges. The simulation modeling of the repair process needs to consider two aspects: the simulation modeling of the repair entity and the simulation modeling of repair strategies.

3.3.1. Repair entity simulation

Repair entity simulation usually consists of five parts: the simulation of the repair team's original position, fault location speed, repair speed, movement strategy, and movement speed [21,22,44].

The original location of repair teams is designed at the node's location. That is, at least one repair team is stationed at the location of each backbone network node. The main consideration is that the location of branches of the network operator is basically the same as the location of the backbone node (computer room, tower station and other core equipment), both in big cities with dense population and convenient transportation. The operators' location selection method can effectively control their operating costs, and it also brings convenience to our simulation modeling. The fault location speed refers to the speed at which the maintenance worker operates the locator to measure the damaged position of the optical cable.

Repair speed is divided into two categories: node repair speed and link repair speed. Node repair generally refers to the repair of damaged optical switches, and link repair generally refers to the fusion of several optical fibers in the optical cable. The repair speed is mainly affected by the number of repair personnel and equipment, the technical proficiency of repair personnel, and repair resources.

Movement strategy and movement speed refer to the travel strategy and average travel speed of the maintenance team's vehicles. Since the network is modeled on a two-dimensional plane, the maintenance team's movement strategy is simplified as this: starting from the node position, the target is the damaged part of the fiber connected to the node, and the direction is to drive straight along with the optical fiber. The symbol definition, calculation formula and value range of fault location speed, node repair speed, link repair speed, and vehicle moving speed will all be given in Section 4.

3.3.2. Repair strategy simulation of SRS

Repair strategy refers to a plan to sequentially repair multiple network damaged elements based on existing maintenance resources. Based on the Saturated Rescue Strategy, in the face of a large-scale disaster in a certain area, the government will coordinate and mobilize all local and surrounding resources to ensure the continuous and adequate supply of maintenance personnel, maintenance equipment and maintenance resources to support the disaster-affected area. The purpose is to complete the emergency repair work of the communication network at the fastest speed to ensure rescue communications and basic communications of the people. Therefore, as maintenance personnel, equipment and maintenance resources continue to be sufficient, the average maintenance speed in SRS can be regarded as unchanged. There are two differences between SRS and resource-constrained rescue strategies. First, there are sufficient rescue personnel and equipment, and the repair of multiple damaged links is carried out at the same time, and there is no need to choose the best. The second is that the average rescue speed remains the same and will not decrease over time. Therefore, SRS allows model users to no longer need to input additional repair resource-constrained data, and the model no longer needs to perform additional optimization calculations, which improves the application scope of the overall model. If it is necessary to simulate resource-constrained conditions, SRS can also respond: model users need to adjust the average repair speed of the specified link according to the actual requirements (the repair speed of all links is the same by default). According to Formula (2), since we set the index number of the link for each virtual point, the model user can set the corresponding judgment sentence according to the link number.

In SRS, repair resources are divided into Node-maintenance Vehicle Groups (NVG) and Link-maintenance Vehicle Groups (LVG). There is one NVG and w_i LVG deployed at each node's location, where w_i is the node connection degree of v_i. Damaged nodes or damaged links are only repaired by one NVG or one LVG. The LVG dispatch rules are as follows: suppose link e_{12} is damaged, v_1 and v_2 are the nodes at both ends of e_{12}, let l_1 be the longest distance from v_1 to the damaged position on e_{12}, and l_2 is the longest distance from v_2 to the damaged position. If there is $l_1 \geq l_2$, LVG on v_2 will be dispatched to repair, otherwise LVG on v_1 will be dispatched. Fig. 3 uses the simple network in Fig. 1 as an example to explain the SRS scheduling rules. In Fig. 3(a), there are two damage circles with different radius that cause damage to the links e_{12}, e_{23}, and e_{14}. l_1 and l_2 are the distances from node v_1 and v_2 to the farthest damaged location of e_{12}, l_3 and l_4 are the distances from node v_2 and v_3 to the farthest damaged location of e_{23}, and l_5 and l_6 are the distances from node v_1 and v_4 to the farthest damaged location of e_{14}. There are $l_1 > l_2$, $l_4 > l_3$, $l_6 > l_5$, therefore LVGs on v_1 and v_2 were dispatched to repair e_{14}, e_{12} and e_{23}. NVG on v_2 was dispatched to repair v_2. In Fig. 3(b), there are $l_1 > l_2$, $l_3 > l_4$, $l_6 > l_5$, $l_7 > l_8$, therefore LVGs on v_2, v_3, v_4 and NVG on v_4 were dispatched.

4. Modeling for network vulnerability analysis

This section details the design principles of the vulnerability analysis model: Damage Measurement and Location Model (DMLM), and details the three new vulnerability analysis metrics. We also propose the value ranges of significant parameters that affect the calculation accuracy of DMLM under different metrics and perform analytical proofs.

4.1. The damage measurement and location model

Since DMLM needs to attack the network first, we designed the Optimized Grid-partition-based Method (OGPM) which is a traversal algorithm based on random points, to select candidate attack locations. The grid-partition-based method is firstly proposed by Wang to accurately locate the damage circle center [5]. It firstly employs a grid to evenly partition the network plane Z into many square grid cells. The grid's diameter is denoted as d. Then randomly select a candidate point b in each grid to form the candidate points set B. When d is small enough, the difference between b and other positions in the grid can be accepted within a specific error range. That is, b is the representation of its grid. Traverse all candidate points as the damage center to attack the network, and the candidate point that caused the worst failure is the best attack position. Fig. 4(a) illustrates an example of the grid-partition-based method. As the figure shows, the plane Z is required to be the smallest rectangle covering the network, and its long side and wide side are parallel to the x-axis and y-axis, respectively.

The original algorithm needs to traverse all candidate points in the network plane, which leads to a large computation complexity. Because of the sparse distribution of actual optical fibers, many traversals of candidate points are in vain. To optimize the method, we exclude these meaningless points by employing geometric calculations with three rectangles O, C, R and their corresponding matrices.

4.1.1. The original rectangle

As shown in Fig. 4(a), the Original Rectangle O is generated firstly. Four edges of O are respectively parallel to the edges of Z, and the distance between each set of two paralleled edges is $2r$. The increased distance $2r$ is to ensure that all grids that may affect the network are included. The lower left vertex of O coincides with the origin of coordinates. Grid O by the grid with diameter d, and get the $N \times M$ grids.

$$N = \mathbf{Int}\left(\frac{Wth + 4r}{d}\right), M = \mathbf{Int}\left(\frac{Lth + 4r}{d}\right). \tag{8}$$

where Wth and Lth are the length of the long side and wide side of Z. Randomly select a point in each grid and form two $N \times M$ matrices: X and Y, which are matrices of x and y coordinates of these random points, respectively. X and Y are denoted as the Original Matrix.

4.1.2. The cover rectangle

The geometric calculation method is illustrated in conjunction with the example in Fig. 4(b). To link e_{12}, the Cover Rectangle C is generated using the following rules. Its long side and wide side are parallel and perpendicular to e_{12}, respectively. The length of the long side is $l_{12} + 2r$ and the length of the wide side is $2r$. The shortest distance from the four sides of C to the edge e_{12} is r. If now attacking link e_{12} with candidate points as circle centers and r as radius, points that could damage e_{12} and its nodes v_1, v_2 will not fall outside C. Four vertexes' coordinates of C can be easily obtained by employing geometric calculation on v_1, v_2, which denoted as (x_1^c, y_1^c), (x_2^c, y_2^c), (x_3^c, y_3^c), (x_4^c, y_4^c).

4.1.3. The range rectangle

The Range Rectangle R is shown in Fig. 4(b). The main function of R is to reduce the matrix elements that need to be traversed when finding candidate points inside the Cover Rectangle. The vertex coordinate formulas of R are affected by the slope K of link e_{12} and can be divided into the following two cases:

$$\begin{aligned}
K \geq 0, \quad & \begin{cases} x_1^R = x_1^C \\ y_1^R = y_2^C \end{cases} \begin{cases} x_2^R = x_3^C \\ y_2^R = y_2^C \end{cases} \begin{cases} x_3^R = x_3^C \\ y_3^R = y_4^C \end{cases} \begin{cases} x_4^R = x_1^C \\ y_4^R = y_4^C, \end{cases} \\
K < 0, \quad & \begin{cases} x_1^R = x_4^C \\ y_1^R = y_1^C \end{cases} \begin{cases} x_2^R = x_2^C \\ y_2^R = y_1^C \end{cases} \begin{cases} x_3^R = x_2^C \\ y_3^R = y_3^C \end{cases} \begin{cases} x_4^R = x_1^C \\ y_4^R = y_3^C. \end{cases}
\end{aligned} \tag{9}$$

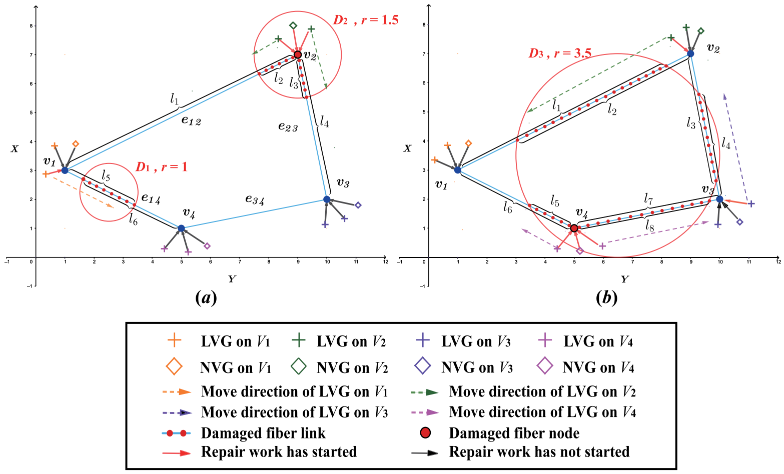

Fig. 3. Examples of SRS scheduling rules. The four-color cross and diamond respectively represent NVG and LVG at the four node positions. The red solid arrow indicates that the NVG or LVG has started repair work, and the different colored dashed arrow indicates the moving direction of the LVG in different position.

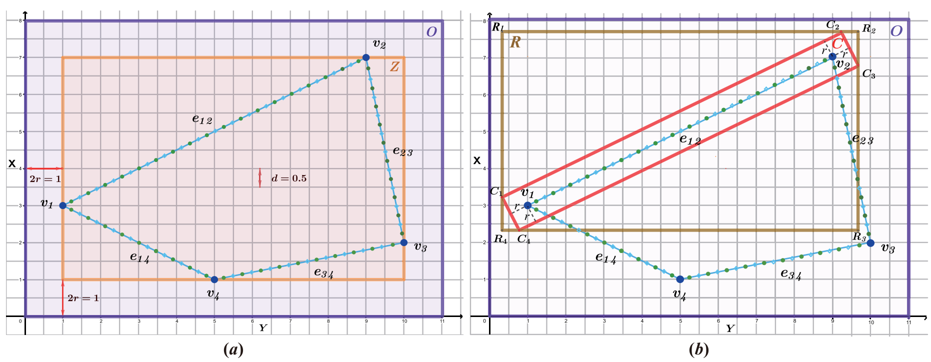

Fig. 4. (a) The grid-partitioned example network and its network plane Z and the Original Rectangle O. (b) The Cover Rectangle C and the Range Rectangle R of the example network.

Two new matrices X', Y' are denoted as the Range Matrix, which can be calculated as follows:

$$X' = X\left(\alpha_1 : \alpha_2, \beta_1 : \beta_2\right), Y' = Y\left(\alpha_1 : \alpha_2, \beta_1 : \beta_2\right), \tag{10}$$

where $\alpha_1 = \textbf{Int}\left(x_1^R/d\right)$, $\alpha_2 = \textbf{Int}\left(x_3^R/d\right) + 1$, $\beta_1 = \textbf{Int}\left(y_1^R/d\right)$, $\beta_2 = \textbf{Int}\left(y_3^R/d\right) + 1$. $\left(\alpha_1 : \alpha_2, \beta_1 : \beta_2\right)$ means all elements within rows α_1 to α_2 and columns β_1 to β_2 of the matrix. Thus, the Cover Matrix X'', Y'' can be calculated as follows:

$$\left[X'', Y''\right] = \textbf{Inrectangle}\left(\left[X', Y'\right], C\right), \tag{11}$$

where the function **Inrectangle** is to find the points located within the given rectangle. X'', Y'' contain the coordinates x, y of all candidate points that may damage e_{12}. It should be noted that there are still a small number of candidate points nearby the four vertexes' positions of C, which cannot damage e_{12}. Considering the computational complexity, we have not excluded these points. The candidate points set of the edge e_{12} is obtained by combining elements at the same position in X'' and Y''. Traverse each edge to obtain its corresponding candidate point sets, summarize these sets and delete the duplicate points, then get the final candidate points set B. The pseudo-code of the optimization algorithm is shown in Algorithm 1. Since the algorithms in our paper involve large-scale matrix operations, the pseudo-codes given are all based on MATLAB.

Algorithm 1 greatly reduces the number of candidate points that need to be traversed. It is independent of other algorithms in DMLM and only needs to be called once initially, thereby greatly improving DMLM's operating efficiency. Algorithm 2 describes the calling sequence of algorithms and models in DMLM. In the algorithm, b_k is

Algorithm 1 The OGPM Algorithm

Input: the topology of network $G(V, E)$, the radius of damage circle r, the grid's diameter d;
Output: the candidate points set B;
$B \leftarrow \varnothing$;
Calculate Wth, Lth of rectangle Z;
Generate X, Y according to Formula (8);
for each link $e_{ij} \in E$ **do**
  Calculate Four vertex coordinates of rectangle C;
  Calculate Four vertex coordinates of rectangle R according Formula (9);
  Generate X', Y' according Formula (10);
  Generate X'', Y'' according Formula (11);
  Combine X'', Y'' and get B_i;
  Delete $B \cap B_i$ in B_i;
  $B \leftarrow B_i \cup B$;
end for
return B

the candidate point calculated by Algorithm 1, and there is $b_k \in B$. Δ is the metric value of the network performance after attacking network on the position of b_k. Δ^* is the metic value of the repair result.

4.2. The Remain Transmitting Capacity Ratio

The RCTR is mainly used for comparison with the newly designed two vulnerability analysis indicators based on network recoverability.

Algorithm 2 The DMLM Algorithm

Input: $G(V, E)$, r, the damage efficiency function D, the grid's diameter d, the segmentation distance s, the metric Δ, the repair parameters

Output: the sorted candidate points set B^*

Call **Algorithm 1**, get B;
for each $b_k \in B$ **do**
  Find v_i, e_{ij} touched by b_k as the center and r as the radius;
  Call the **LSM** on these e_{ij}, get virtual nodes p^{ij} with the segmentation distance s;
  Call the **DECM** on these v_i and p^{ij}, get their damage degree: Δ_k;
  Start the repair process with the **SRS** on the damaged v_i and p^{ij}, get Δ_k^* based on Δ_k;
end for
Combine B with corresponding Δ_k^*, get B^*;
Sort B^* in descending or ascending order according to Δ_k^*;
return B^*

The design of RCTR imitates the metrics: the Remaining Link Capacity [5] and the Total Remaining Link Capacity [11]. Both metrics describe the remaining transmission capacity of links after damage. RTCR is a statistical metric that measures the damage degree to network equipment after the geographical failure. It records the ratio between the transmitting capacity F of all equipment after a disaster or attack and the initial transmitting capacity. RCTR calculation does not need to start the repair process. Considering the difference in the equipment's invulnerability of nodes and links, we separately calculate the loss ratio of the two.

4.2.1. The linear function with RCTR

We firstly discuss the linear function. According to Formula (7), the calculating equation for the node's RCTR is proposed as follows.

$$\gamma_{node} = \frac{\sum_{i=1}^{n_v} \left(F_i - (1-\varphi_v) \cdot D(g_i)\right)}{\sum_{i=1}^{n_v} F_i}, \tag{12}$$

where g_i is the distance of v_i to the damage circle center, and n_v is the number of nodes. Substituting Formula (5) into the above formula, thereby it can be transformed as follows:

$$\gamma_{node} = \frac{\left(\sum^{n_v} F_i - n_d\right) + \varphi_v \cdot n_d + a\left(1-\varphi_v\right) \cdot \sum_i^{n_d} g_i/r}{\sum^{n_v} F_i}, \tag{13}$$

where n_d is the number of damaged nodes. The following theorem expresses how to select a proper grid diameter d to satisfy the $(1-\varepsilon_1)$-approximation to the node's RCTR.

Theorem 1. *The grid partition-based method based on the linear function induces the $(1-\varepsilon_1)$-approximation to the node's RCTR if the grid diameter d satisfies the following condition:*

$$d \le \frac{r \cdot \varepsilon_1 \cdot \varphi_v}{\sqrt{2}a\left(1-\varepsilon_1\right)\left(1-\varphi_v\right)}. \tag{14}$$

Proof. The relative position of the node v_1 and a grid square z_1 is shown in Fig. 5. To meet the $(1-\varepsilon_1)$-approximation requirements of the metric results in Theorem 1, the metric results ratio of any two candidate points b_x, b_y in the same grid z_1 must meet the error requirements as follows:

$$min\left\{\frac{\gamma_{node}(b_x)}{\gamma_{node}(b_y)}, \frac{\gamma_{node}(b_y)}{\gamma_{node}(b_x)}\right\} \ge 1-\varepsilon_1. \tag{15}$$

Assuming that γ_{node} takes the minimum and maximum values at b_x and b_y respectively, then there is

$$min\left\{\frac{\gamma_{node}(b_x)}{\gamma_{node}(b_y)}, \frac{\gamma_{node}(b_y)}{\gamma_{node}(b_x)}\right\} \ge \frac{\gamma_{node}^{-}(b_x)}{\gamma_{node}^{+}(b_y)} \ge 1-\varepsilon_1. \tag{16}$$

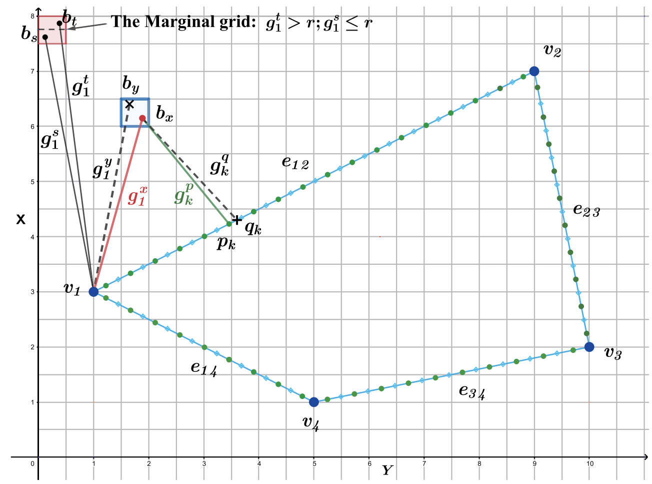

Fig. 5. Examples of changes in candidate points' position and virtual points' position.

Put Formula (13) into (16), there is

$$\frac{r\left(\sum_i^{n_v} F_i - n_d\right) + n_d \cdot r \cdot \varphi_v + a\left(1-\varphi_v\right)\sum_i^{n_d} g_i^x}{r\left(\sum_i^{n_v} F_i - n_d\right) + n_d \cdot r \cdot \varphi_v + a\left(1-\varphi_v\right)\sum_i^{n_d} g_i^y} \ge 1-\varepsilon_1, \tag{17}$$

where g_i^x, g_i^y express the distance of v_i to b_x, b_y. As shown in Fig. 5, no matter how the relative position of v_i and z_1 change, there is $g_i^y \le g_i^x + \sqrt{2}d$. Without loss of generality, there is

$$\frac{r\left(\sum_i^{n_v} F_i - n_d\right) + n_d \cdot r \cdot \varphi_v + a\left(1-\varphi_v\right)\sum_i^{n_d} g_i^x}{r\left(\sum_i^{n_v} F_i - n_d\right) + n_d \cdot r \cdot \varphi_v + a\left(1-\varphi_v\right)\sum_i^{n_d}\left(g_i^x + \sqrt{2}d\right)} \ge 1-\varepsilon_1. \tag{18}$$

Sort out the formula, there is

$$d \le \frac{\varepsilon_1 \cdot \varphi_v \cdot r \cdot n_d + a \cdot \varepsilon_1\left(1-\varphi_v\right) \cdot \sum_i^{n_d} g_i^x + \varepsilon_1 \cdot r\left(\sum_i^{n_v} F_i - n_d\right)}{\sqrt{2}a \cdot n_d\left(1-\varepsilon_1\right)\left(1-\varphi_v\right)}. \tag{19}$$

Because $g_i^x \ge 0$ there is $\sum_i^{n_d} g_i^x \ge 0$. The equal sign is established when $n_v = 1$ and b_x coincide with v_i. When the fiber backbone network meets any two of the following conditions, there is $n_d \le \sum_i^{n_v} F_i$.

Condition (i): the transmission capacity of each device in the optical fiber network is similar. Most of the actual operating fiber backbone networks meet this condition. In a normally operating fiber backbone network, similar transmission capabilities between devices are a necessary condition to avoid network congestion. The service flow transmitted in the network can be dynamically adjusted according to the needs of users, but the fiber backbone network as the support network does not allow an excessive gap in transmission capacity. In ideal conditions, there is $\forall F_i \approx 1$. Thus, $\sum_i^{n_v} F_i \approx n_v \ge n_d$.

Condition (ii): the damage coverage (the damage circle area) is much smaller than the coverage of the fiber backbone network on the two-dimensional plane. This condition ensures that the number of damaged devices on the network is much smaller than the number of all devices on the network. Since the coverage of the fiber backbone network is much larger than the effective impact range of common disasters, most damage scenarios meet Condition (ii). Thus, $n_v > \sum_i^{n_v} F_i \ge n_d$.

When the fiber backbone network meets Condition (i) or (ii), there is $n_d \le \sum_i^{n_v} F_i$. Without loss of generality, there is $\sum_i^{n_d} g_i^x = 0$, then condition (14) can be deduced.

Besides, a special case needs to be discussed: the marginal grid. As shown in Fig. 5, in the marginal grid, the distance from some positions to a node is larger than r, and the distance from the remaining positions to the node is less than or equal to r. Since the number of marginal grids is small and in the worst position, their influence on the calculation of the optimal solution is very limited. Therefore, this paper does not consider the special case of the marginal grid when discussing the value range of d and s.

For the link's RCTR, F of the entire link can be obtained by accumulating F of the sub-line segments, and F of the sub-line segment can be obtained by integrating F of the virtual point with the length s. Thereby, the formula is proposed as follows:

$$\gamma_{link} = \frac{\sum_{q=1}^{n_e} \sum_{k=1}^{m_q} \int_0^s \left(F_k^q - (1-\varphi_e) \cdot D\left(g_k^q\right)\right)}{\sum_{q=1}^{n_e} \sum_{k=1}^{m_q} \int_0^s F_k^q}, \tag{20}$$

where q is the index value of link sort, n_e is the number of links, m_q is the number qth link's line segments, and k is the index value of line segment sort. We sort out the formula and calculate the integral, then there is

$$\gamma_{link} = \frac{\sum_{k=1}^{M_v} \left(F_k - (1-\varphi_e) \cdot D(g_k)\right) \cdot s}{\sum_{k=1}^{M_v} F_k \cdot s} = \frac{\sum_{k=1}^{M_v} \left(F_k - (1-\varphi_e) \cdot D(g_k)\right)}{\sum_{k=1}^{M_v} F_k}, \tag{21}$$

where M_v is the number of all virtual points. Obviously, Formula (21) is very similar to Formula (12). Thus, we propose Lemma 1 to select a proper d to satisfy the $(1-\varepsilon_1)$-approximation to the link's RCTR.

Lemma 1. *The grid partition-based method based on the linear function induces the $(1-\varepsilon_1)$-approximation to the link's RCTR if d satisfies the following condition:*

$$d \le \frac{r \cdot \varepsilon_1 \cdot \varphi_e}{\sqrt{2}a\left(1-\varepsilon_1\right)\left(1-\varphi_e\right)}. \tag{22}$$

Proof. The proof process is consistent with the node's RCTR.

In the following formula, we perform weighted calculations on γ_{node} and γ_{link} to obtain the γ of the whole network,

$$\gamma = \lambda \cdot \gamma_{node} + (1-\lambda) \cdot \gamma_{link}, \tag{23}$$

where $\lambda \in [0,1]$, and it is adjusted based on the relative importance of nodes and links in the network. Because $\varphi_v > \varphi_e$, Formula (22) is employed to determine the value range of d.

Since we employed a large number of line segments to approximate the simulation results of links, it needs to properly select the segmentation distance s to satisfy the $(1-\varepsilon_2)$-approximation to the link's RCTR. Theorem 2 defines the suitable range of s.

Theorem 2. *The Link Segmentation Model based on the linear function induces the $(1-\varepsilon_2)$-approximation to the link's RCTR if s satisfies the following condition:*

$$s \le \frac{2 \cdot r \cdot \varepsilon_2 \cdot \varphi_e}{a\left(1-\varepsilon_2\right)\left(1-\varphi_e\right)}. \tag{24}$$

Proof. as Fig. 5 shows, we assume the candidate point locates at b_x. Because the virtual point p_k is the midpoint of the line segment, the distance between b_x and p_k is g_k^p. Assume there is a random point q_k near p_k in the same segment, its distance to b_x is g_k^q. To meet the $(1-\varepsilon_2)$-approximation requirement in Theorem 2, the metric results ratio must meet the error requirements as follows:

$$min\left\{\frac{\gamma_{link}(p_1,\dots,p_k)}{\gamma_{link}(q_1,\dots,q_k)}, \frac{\gamma_{link}(q_1,\dots,q_k)}{\gamma_{link}(p_1,\dots,p_k)}\right\} \ge 1-\varepsilon_2. \tag{25}$$

Assume there is $\sum_{k=1}^{M_v} g_k^p \ge \sum_{k=1}^{M_v} g_k^q$, M_v is the number of virtual points. Formula (25) is transformed as follows:

$$\frac{r\left(\sum_i^{M_v} F_i - M_d\right) + M_d \cdot r \cdot \varphi_e + a\left(1-\varphi_e\right)\sum_{k=1}^{M_d} g_k^q}{r\left(\sum_i^{M_v} F_i - M_d\right) + M_d \cdot r \cdot \varphi_e + a\left(1-\varphi_e\right)\sum_{k=1}^{M_d} g_k^p} \ge 1-\varepsilon_2. \tag{26}$$

As shown in Fig. 5, there is $g_k^p \le g_k^q + s/2$. Substitute their relationship into the calculation, then there is

$$s \le \frac{2\varepsilon_2 \cdot \varphi_e \cdot r \cdot M_d + 2a \cdot \varepsilon_2\left(1-\varphi_e\right) \cdot \sum_{k=1}^{M_d} g_k^q + 2\varepsilon_2 \cdot r\left(\sum_i^{M_v} F_i - M_d\right)}{aM_d\left(1-\varepsilon_2\right)\left(1-\varphi_e\right)}. \tag{27}$$

When the fiber backbone network meets Condition (i) or Condition (ii), there is $\sum_i^{M_v} F_i \ge M_d$, and because of $\sum_{k=1}^{M_d} g_k^q \ge 0$, condition (24) is deduced. Similarly, when $\sum_{k=1}^{M_v} g_k^p \le \sum_{k=1}^{M_v} g_k^q$, condition (24) also holds.

It should be noted that in actual calculations, if the r is large, the calculated s may also be large. However, we recommend that s does not exceed 2 kilometers. The reason is that the current production length of optical cables is generally 1 ∼ 2 kilometers, that is, every 1 ∼ 2 kilometers requires an adapter box to splice two sections of fiber cable. The junction is more vulnerable to damage, so setting the s to be less than 2 kilometers has practical significance for the reliability analysis of the optical cable. Model users can also set the value of s according to the requirements of damage simulation in the experiment or the decision-making content resulted from the vulnerability analysis.

4.2.2. The Gaussian distribution function with RCTR

We first discuss the Gaussian distribution function with the node's RCTR.

Theorem 3. *The grid partition-based method based on the Gaussian distribution function induces the $(1-\varepsilon_1)$-approximation to the node's RCTR if the grid diameter d satisfies the following condition:*

$$d \le \frac{r}{\sqrt{2\omega}} \ln^{\frac{1}{2}}\left(1 + \frac{\varepsilon_1 \varphi_v}{1-\varepsilon_1-\varphi_v}\right). \tag{28}$$

Proof. Substituting Formula (6) into Formula (15) gives:

$$\frac{\sum_i^{n_v} F_i - (1-\varphi_v)\sum_i^{n_d} e^{-\frac{\omega}{r^2}(g_i^x)^2}}{\sum_i^{n_v} F_i - (1-\varphi_v)\sum_i^{n_d} e^{-\frac{\omega}{r^2}(g_i^y)^2}} \ge 1-\varepsilon_1, \tag{29}$$

According to Fig. 5, there is $g_i^y \le g_i^x + \sqrt{2}d$. Thus Formula (29) is transformed as follows:

$$\begin{aligned} & e^{-\frac{\omega}{r^2}(g_1^x)^2}\left(1-(1-\varepsilon_1)e^{-\frac{2\omega}{r^2}d^2} \cdot e^{-\frac{2\sqrt{2}\omega}{r^2}d \cdot g_1^x}\right) + \cdots \\ & + e^{-\frac{\omega}{r^2}(g_{n_d}^x)^2}\left(1-(1-\varepsilon_1)e^{-\frac{2\omega}{r^2}d^2} \cdot e^{-\frac{2\sqrt{2}\omega}{r^2}d \cdot g_{n_d}^x}\right) \le \frac{\varepsilon_1 \sum_i^{n_v} F_i}{1-\varphi_v}. \end{aligned} \tag{30}$$

Because $\forall e^{-\frac{2\sqrt{2}\omega}{r^2}d \cdot g_i^x} \in (0,1]$, it can be derived from Formula (30):

$$\sum_i^{n_d} e^{-\frac{\omega}{r^2}(g_i^x)^2}\left(1-(1-\varepsilon_1)e^{-\frac{2\omega}{r^2}d^2}\right) \le \frac{\varepsilon_1 \sum_i^{n_v} F_i}{1-\varphi_v}. \tag{31}$$

Sorting Formula (31), and the natural logarithm is taken on both sides of the inequality to obtain:

$$d^2 \le -\frac{r^2}{2\omega} \ln\left(\frac{1}{1-\varepsilon_1} - \frac{\varepsilon_1 \sum_i^{n_v} F_i}{(1-\varepsilon_1)(1-\varphi_v)\sum_i^{n_d} e^{-\frac{\omega}{r^2}(g_i^x)^2}}\right). \tag{32}$$

There is $\sum_i^{n_d} e^{-\frac{\omega}{r^2}(g_i^x)^2} \le n_d$. When fiber backbone meets the two above conditions, there is $\sum_i^{n_d} e^{-\frac{\omega}{r^2}(g_i^x)^2} \le n_d \le \sum_i^{n_v} F_i$. Adjust the formula and sqrt the inequality to get condition (28).

We then give the Gaussian distribution function with the edge's RCTR.

Lemma 2. *The grid partition-based method based on the Gaussian distribution function induces the $(1-\varepsilon_1)$-approximation to the edge's RCTR if the grid diameter d satisfies the following condition:*

$$d \le \frac{r}{\sqrt{2\omega}} \ln^{\frac{1}{2}}\left(1 + \frac{\varepsilon_1 \varphi_e}{1-\varepsilon_1-\varphi_e}\right). \tag{33}$$

Proof. According to Formula (21), in the ideal condition, the transmission capacity of each link is the same, the proof process would be consistent with the node's RCTR. The calculation method of network's γ is the same as Formula (23), and will not be repeated.

It also needs to properly select the segmentation distance s based on the Gaussian distribution function to satisfy the $(1-\varepsilon_2)$-approximation, Theorem 4 defines the suitable range of s.

Theorem 4. *The Link Segmentation Model based on the linear function induces the* $(1-\varepsilon_2)$*-approximation to the link's RCTR if* s *satisfies the following condition:*

$$s \leq \frac{2r}{\sqrt{\omega}} \ln^{\frac{1}{2}} \left(1 + \frac{\varepsilon_2 \varphi_e}{1-\varepsilon_2-\varphi_e}\right). \tag{34}$$

Proof. According to Formula (25), we assume there is $\sum_{k=1}^{M_v} g_k^p \geq \sum_{k=1}^{M_v} g_k^q$. According to Fig. 5, there is $g_k^p \leq g_k^q + s/2$. Substituting Formula (6) into Formula (25) gives:

$$\frac{\sum_i^{M_v} F_i - (1-\varphi_e)\sum_i^{M_d} e^{-\frac{\omega}{r^2}(g_k^q)^2}}{\sum_i^{M_v} F_i - (1-\varphi_e)\sum_i^{M_d} e^{-\frac{\omega}{r^2}(g_k^q+s/2)^2}} \geq 1-\varepsilon_2, \tag{35}$$

where M_v is the number of virtual points, and M_d is the number of damaged virtual nodes. Sorting the above formula and obtain:

$$\sum_i^{M_d} e^{-\frac{\omega}{r^2}(g_k^q)^2}\left(1-(1-\varepsilon_2)e^{-\frac{\omega}{4r^2}s^2}\right) \leq \frac{\varepsilon_2\sum_i^{M_v} F_i}{1-\varphi_e}. \tag{36}$$

Arranging the inequality and taking the limit, condition (34) could be deduced. Similarly, when $\sum_{k=1}^{M_v} g_k^p \leq \sum_{k=1}^{M_v} g_k^q$, condition (34) also holds.

It should be pointed out that in some special experimental scenarios, the optical fiber network may not meet the above Condition(i)(ii). Such as the process of upgrading and updating network equipment, there may be an excessive gap in the transmission capabilities of the backbone network. Or when experimenters conduct extreme damage tests on the optical fiber network (the damage circle covers almost the entire optical fiber network). We recommend dividing networks with too large gaps into different sub-networks for vulnerability analysis. Otherwise, high-risk areas in the network will be concentrated near facilities with higher transmission capacity, which will affect the accuracy of the vulnerability analysis results.

4.3. The Maximum Repair Time (MRT)

The network repair time is mainly affected by the various parameters of the repair entity, the repair strategy and the geographical distribution of network elements. We combine DMLM and SRS to give calculation formulas for node and link repair time. The following formula describes the node repair time.

$$t_i^v = (1-\varphi_v)\cdot D(g_i)/\mu_v, \tag{37}$$

where t_i^v is the repair time of v_i, and μ_v is the node repair velocity, which indicates how much transmitting capacity is repaired per unit time. There are many types of damage to the nodes of the optical fiber network. The most common occurrence is the bending or desoldering of the optical switch motherboard under physical impact. For this type of damage, a single skilled worker can complete the repair of a single machine in 30–40 min. The specific value of μ_v is determined according to the number of optical switches and the number of workers in the computer room, and we will give it in detail in Section 5 in conjunction with the experiment.

The link repair time is proposed as follows:

$$t_{ij}^e = t_{repair} + t_{travel} + t_{locate} = \frac{\sum_{k=1}^{M_d}(1-\varphi_e)\cdot D(g_k)}{\mu_s} + \frac{l_d+l_t}{\mu_t} + \frac{l_d}{\mu_l}, \tag{38}$$

where t_{ij}^e is the repair time of e_{ij}. It consists of three parts: the damage repair time t_{repair}, the LVG traveling time t_{travel}, the fault location time t_{locate}. In t_{repair}, M_d is the number of damaged virtual points, and μ_s indicates how much transmitting capacity of the line segment is repaired per unit time. The line segment's length s is inversely proportional to the repair transmitting capacity per unit time on the line segment. Thus, there is $\mu_s = \mu_e/s$, where μ_e indicates how much transmitting capacity of the link with unit distance is repaired per unit time. Common fiber damage is bending, disconnection or poor contact in the connector box. The repair method is fiber splicing or splice replacement. A single skilled worker can splice a completely broken fiber in 3 min. The specific value of μ_e is determined according to the number of fibers in the optical cable and the number of workers. In t_{travel}, l_d is the length of the damaged part in the link, l_t is the distance from the LVG to the nearest damaged position from the node, and μ_t is the vehicle speed. In t_{locate}, μ_l indicates the length of the damaged link can be located per unit time. Using the optical time-domain reflectmeter (OTDR) to locate a single fault point with an accuracy of 10 meters takes 6 min. The values of μ_e, μ_t, and μ_l will also be specifically given in Section 5 in conjunction with the experiment. The MRT T is proposed as follows:

$$T = max\left\{t_1^v,\dots,t_i^v,t_{12}^e,\dots,t_{ij}^e\right\}. \tag{39}$$

In the following theorems, we express the way to select a proper grid d to satisfy the $(1-\varepsilon_1)$-approximation to the node and link's MRT.

4.3.1. The linear function with MRT

We first discuss the node's MRT based on the linear function.

Theorem 5. *The grid partition-based method induces the* $(1-\varepsilon_1)$*-approximation to the node's MRT if* d *satisfies the following condition:*

$$d \leq \varepsilon_1 \cdot r \cdot (1-a)/\sqrt{2}a. \tag{40}$$

Proof. Similar to the above proof process, according to Formula (37), the ratio of metric results of any two candidate points b_x, b_y in the same grid z_1 must meet the error requirements as follows:

$$\frac{r - a\cdot g_i^y}{r - a\cdot g_i^x} \geq 1-\varepsilon_1. \tag{41}$$

Because there is $g_i^y \leq g_i^x + \sqrt{2}d$, then Formula (41) can be transformed as $d \leq (\varepsilon_1\cdot r - a\cdot\varepsilon_1\cdot g_i^x)/\sqrt{2}a$. Because $g_i^x < r$, and without loss of generality, condition (40) holds. We then discuss the edge's MRT based on the linear function.

Lemma 3. *The grid partition-based method induces the* $(1-\varepsilon_1)$*-approximation to the edge's MRT, if* d *satisfies the following condition:*

$$d \leq min\left\{r\varepsilon_1(1-a)/\sqrt{2}a, r\sqrt{\varepsilon_1(1-\varepsilon_1)}\right\}. \tag{42}$$

Proof. Since t_{ij}^e is composed of three parts, each part needs to induce the value range of d separately. For t_{travel}, its proof is similar with the proof of Formula (40). Substituting the first part of Formula (38) into the ratio of metric results of any two candidate points b_x, b_y and obtain:

$$\frac{rM_d - a\sum_i^{M_d} g_i^y}{rM_d - a\sum_i^{M_d} g_i^x} \geq 1-\varepsilon_1. \tag{43}$$

Because $g_i^y \leq g_i^x + \sqrt{2}d$, Formula (43) could be arranged as follows:

$$d \leq \frac{r\varepsilon_1\left(M_d - \frac{a}{r}\sum^{M_d} g_i^x\right)}{\sqrt{2}aM_d}. \tag{44}$$

Since $\sum^{M_d} g_i^x \leq rM_d$, the left part of condition (42) can be deduced. For t_{travel}, the distance between the grid and the link will seriously affect

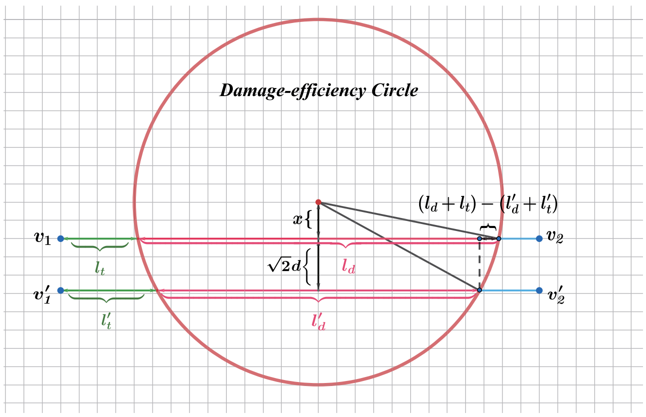

Fig. 6. Example of calculation of link's MRT.

the value range of d. In general, the closer the grid is to the link, the greater the maximum value of d that can be obtained. We only discuss the case which b moves in the vertical direction of the link. As shown in Fig. 6, l_t' and l_d' are the new l_t and l_d formed after the damage circle moves vertically. The error approximation method can be designed as follows:

$$\frac{\left(l_t'+l_d'\right)/\mu_t}{\left(l_t+l_d\right)/\mu_t}=\frac{\left(l_t+l_d-\sqrt{r^2-x^2}+\sqrt{r^2-\left(x+\sqrt{2}d\right)^2}\right)/\mu_t}{\left(l_t+l_d\right)/\mu_t}\geq 1-\varepsilon_1. \tag{45}$$

Sort out (45) and get

$$d\leq\frac{\sqrt{2}}{2}\cdot\left(\sqrt{r^2-\left(\sqrt{r^2-x^2}-\varepsilon_1\left(l_t+l_d\right)\right)^2}-x\right)\quad x\in\left[0,r-\sqrt{2}d\right]. \tag{46}$$

We assume that $r=5$, $\varepsilon_1=0.1$. If $x=0$, $l_d+l_t=11$, there is $d\leq 2.21$, and if $x=r-\sqrt{2}d$, $l_d+l_t=1$, there is $d\leq 7\times10^{-4}$. d differs greatly under different conditions of x and l_d+l_t, which leads to the inability to induces the $(1-\varepsilon_1)$-approximation to t_{travel}, as well as t_{locate}. But it should be noted that when the candidate point b is in a more effective attack position, i.e. when $x\to0$ and $l_d\to 2r$, there is $d\leq r\sqrt{\varepsilon_1(1-\varepsilon_1)}$. In the same way, d of t_{locate} can also be calculated by this method, as $d\leq r\sqrt{\varepsilon_1(1-\varepsilon_1/2)}$. Due to $\sqrt{\varepsilon_1(1-\varepsilon_1)}<\sqrt{\varepsilon_1(1-\varepsilon_1/2)}$, therefore, condition (42) can be deduced.

Theorem 6. *The Link Segmentation Model based on the linear function induces the $(1-\varepsilon_2)$-approximation to the link's MRT if s satisfies the following condition:*

$$s\leq 2r\varepsilon_2(1-a)/a \tag{47}$$

Proof. The metric results ratio must meet the error requirements as follows:

$$\frac{(1-\varphi_e)\left(M_d-a/r\sum_{k=1}^{M_d}g_k^q\right)/\mu_s+\left(l_t+l_d\right)/\mu_t+l_d/\mu_l}{(1-\varphi_e)\left(M_d-a/r\sum_{k=1}^{M_d}g_k^p\right)/\mu_s+\left(l_t+l_d\right)/\mu_t+l_d/\mu_l}\geq 1-\varepsilon_2, \tag{48}$$

where M_d is the number of damaged virtual points. There are $g_k^q\leq g_k^p+s/2$ and $\sum_{k=1}^{M_d}g_k^p\leq M_d\cdot r$, then (48) is sorted out as follows:

$$s\leq\frac{2r\varepsilon_2(1-a)}{a}+\frac{2r\varepsilon_2\mu_e}{saM_d(1-\varphi_e)}\left(\frac{l_t+l_d}{\mu_t}+\frac{l_d}{\mu_l}\right). \tag{49}$$

Because there are $l_t\geq0$ and $l_d\geq M_d\cdot s$, when l_t and l_d obtain the minimum value, condition (47) holds.

4.3.2. The Gaussian distribution function with MRT

We first discuss the node's MRT based on the Gaussian distribution function.

Theorem 7. *The grid partition-based method based on the Gaussian distribution function induces the $(1-\varepsilon_1)$-approximation to the node's MRT if the grid diameter d satisfies the following condition:*

$$d\leq\frac{r}{\sqrt{2\omega}}\ln^{\frac{1}{2}}\left(\frac{1}{1-\varepsilon_1}\right). \tag{50}$$

Proof. We directly give the arranged inequality: $e^{-\frac{\omega}{r^2}\left(2d^2+2\sqrt{2}dg_i^x\right)}\geq 1-\varepsilon_1$. Because there is $e^{-\frac{\omega}{r^2}2\sqrt{2}dg_i^x}\leq1$, and take the natural logarithm on both sides of the inequality to get $d^2\leq -r^2/2\omega\ln\left(1-\varepsilon_1\right)$. Condition (50) can be deduced after sqrting the inequality. We then discuss the edge's MRT based on the Gaussian distribution function.

Lemma 4. *The grid partition-based method based on the Gaussian distribution function induces the $(1-\varepsilon_1)$-approximation to the edge's MRT if the grid diameter d satisfies the following condition:*

$$d\leq min\left\{\frac{r}{\sqrt{2\omega}}\ln^{\frac{1}{2}}\left(\frac{1}{1-\varepsilon_1}\right),r\sqrt{\varepsilon_1(1-\varepsilon_1)}\right\}. \tag{51}$$

Proof. Similar to the proof of Formula (35), the arranged inequality is as follows:

$$\frac{\sum^{M_d}e^{-\frac{\omega}{r^2}(g_i^x+\sqrt{2}d)^2}}{\sum^{M_d}e^{-\frac{\omega}{r^2}(g_i^x)^2}}\geq 1-\varepsilon_1, \tag{52}$$

where M_d is the number of damaged virtual points. Formula (52) could be sorted out as follows:

$$\sum^{M_d}e^{-\frac{\omega}{r^2}(g_i^x)^2}\left(e^{-\frac{2\omega}{r^2}d^2}-1+\varepsilon_1\right)\geq 0. \tag{53}$$

Due to $\sum^{M_d}e^{-\frac{\omega}{r^2}(g_i^x)^2}\geq0$, the left part of condition (51) can be deduced. The proof of the right part of condition (51) is completely same with Formula (45).

Theorem 8. *The Link Segmentation Model based on the Gaussian distribution function induces the $(1-\varepsilon_2)$-approximation to the link's MRT if s satisfies the following condition:*

$$s\leq\frac{2r}{\sqrt{\omega}}\ln^{\frac{1}{2}}\left(\frac{1}{1-\varepsilon_2}\right). \tag{54}$$

Proof. The metric results ratio must meet the error requirements as follows:

$$\frac{(1-\varphi_e)\sum^{M_d}e^{-\frac{\omega}{r^2}(g_k^q)^2}/\mu_s+\left(l_t+l_d\right)/\mu_t+l_d/\mu_l}{(1-\varphi_e)\sum^{M_d}e^{-\frac{\omega}{r^2}(g_k^p)^2}/\mu_s+\left(l_t+l_d\right)/\mu_t+l_d/\mu_l}\geq 1-\varepsilon_2. \tag{55}$$

There are $g_k^q\leq g_k^p+s/2$, and let $L=\left(l_t+l_d\right)/\mu_t+l_d/\mu_l$, the inequality can be arranged as:

$$s^2\leq-\frac{4r^2}{\omega}\ln\left|1-\varepsilon_2-\frac{L\varepsilon_2\mu_s}{(1-\varphi_e)\sum e^{-\frac{\omega}{r^2}(g_k^p)^2}}\right|. \tag{56}$$

Due to $L\varepsilon_2\mu_s/(1-\varphi_e)\sum e^{-\frac{\omega}{r^2}(g_k^p)^2}>0$, there is $s^2\leq-\frac{4r^2}{\omega}\ln(1-\varepsilon_2)$. Sqrt the inequality, and condition (54) can be deduced.

4.4. The Continuous Performance Degradation (CPD)

Network recovery is a gradual process. The degradation of network performance, such as the loss of link bandwidth, is not recovered immediately at a certain moment, but continues and gradually reduces over a while. CPD describes the cumulative value of the network performance degradation during the repair process, which is generated by the combination of the repair time of the damaged equipment and the network performance metric. The result of the damaged network's CPD is denoted as $Loss$. Most network performance metrics can be embedded in CPD, such as the average shortest path, the size of the giant component, the natural connectivity [45], etc. We employ the Average Shortest Path (ASP) of the network as the network performance metric in this paper. Because the ASP is calculated based on the network topology graph, the calculation function of ASP is denoted as $ASP(G)$, and its result is denoted as P_{th}. P_{th} is the average of the minimum number of hop count between any two nodes. P_{th} becomes larger when the network suffers damage. If the network collapses into multiple disconnected parts, there is $P_{th} = \infty$. This is inconsistent with the actual situation because many countries can deploy enough microwave communication vehicles or satellite communication vehicles in damaged areas in a short time to temporarily maintain basic communication [6,13,14]. However, no matter what kind of emergency communication vehicle, its transmission delay is far greater than the intact optical cable. Therefore, if a link is damaged and makes $P_{th} = \infty$, we set the hop count of the two nodes of the damaged link as Num, where Num is the number of network nodes. If a node is damaged, its count hops to its neighbors are all Num.

To calculate the network's CPD, we employ a matrix H to record the index values of damaged nodes or edges and their repair time:

$$H_{n\times 3} = \begin{pmatrix} i & i & t^v \\ \vdots & \vdots & \vdots \\ i & j & t^e \end{pmatrix}. \tag{57}$$

Matrix H is arranged in positive order according to the repair time in the third column. Since the first two columns of matrix H are the index values of damaged nodes or edges, so $G(H(1:n,1:2))$ could describe the damaged network topology graph. The performance degradation value when the network is just damaged is $loss_0 = ASP(G(H(1:n,1:2))) - ASP(G)$. The pseudo-code of CPD algorithm is described in Algorithm 3.

Algorithm 3 The CPD Algorithm

Input: $G(V,E)$, $H_{n\times 3}$, function ASP
Output: the damaged network's CPD $Loss$
1: $Loss \leftarrow 0$, $t \leftarrow 0$, $ASP_0 \leftarrow ASP(G)$;
2: **for** $i = 1$ to n **do**
3: $\quad Loss \leftarrow Loss + (ASP(G(H(i:n,1:2))) - ASP_0) \times (H(i,3) - t)$;
4: $\quad t \leftarrow H(i,3)$;
5: **end for**
6: **return** $Loss$

Because the CPD employs the repair time of nodes and links in its algorithm, the value ranges of d and s to satisfy the approximation of the CPD are the same with the MRT.

5. Numerical results

5.1. Model validity analysis

According to the above three metrics, combined with the example topology in Fig. 1, we tested the effectiveness of the LSM, DECM and DMLM. The transmission capacity of each network device has been shown in Fig. 1. Although there is a certain gap in the transmission capacity of each device, we set $r = 1$ to enable the example network to meet the Condition (ii). In the case of determining the damage scenario (r, a and ω), transmission capacity (F and F_p), equipment invulnerability (φ_v and φ_e) and repair conditions (μ_v, μ_e, μ_t, μ_l), the metrics' corresponding results are calculated by adjusting the value of ε_1 and ε_2. We repeat the entire DMLM calculation process 100 times and calculate the average value and the standard deviation of the ratios between the calculated results and the optimal solution. The optimal solution of RCTR and MRT can be obtained through geometric analysis and integral calculation, which are denoted as γ^* and T^*, respectively.

Table 1
The common parameter values in the DMLM in model validity analysis.

r	ω	a	φ_v	φ_e	μ_v (1/h)	μ_e (1/h · km^{-1})	μ_t (km/h)	μ_l (km/h)
1	2	0.6	0.3	0.1	0.5	0.2	30	1

For the γ^* of RCTR, because the sample network only has 4 nodes, the impact of damage on γ_{node} is much greater than γ_{link}. Therefore, even if we set $\lambda = 0.5$, the calculated optimal attack position was still very close to node v_1. For the linear function, the best position is on $(1,3)$. For the Gaussian distribution function, there are two best positions: one is $(1,3)$, the other is near $(1.1,3)$. We take $(1.1,3)$ as the attack position and the integral result is $\gamma = 0.873026$, and the result of $(1,3)$ is $\gamma = 0.873021$. The difference in the result exceeds the calculation accuracy of the model. Due to factual errors in the link simulation calculation, and the length of the damaged link when the attack position is at $(1.1,3)$ is greater than the others. Therefore, the best attack position calculated by the algorithm is near $(1.1,3)$.

For the T^* of MRT, since the edge's repair time is much longer than the node, for the two types of damage function, the optimal attack position should be at the midpoint of the longest edge e_{12}. The coordinates are $(5,5)$.

For CPD, The optimal solution is denoted as $Loss^*$. Although it is a simulation metric, it can still analyze and solve the optimal solution position in the simple example network. For the two types of damage functions, the best attack position will be at the position of the distance r from the node v_1 on the bisector of the angle between e_{12} and e_{14}. The coordinates are $(2,3)$.

Table 1 shows the constant parameter values in experiments. The invulnerability of network equipment is designed based on literature [1,14]. The network repair parameters are set based on literature [6,13,14,46], the fiber repair speed of skilled workers using fiber fusion splicer, and the fault positioning speed of skilled workers using OTDR. We assumed there are 1 skilled worker in the NVG, 10 skilled workers in the LVG, 4 fiber switches in each engine room and 8 fibers in each optical cable. Combining the basic data of each parameter given in Section 4, the values in Table 1 can be calculated. Table 2 (linear function) and Table 3 (Gaussian distribution function) show the average (the left column of the table cell) and standard deviation (the right column) of the ratio between the calculated results and the optimal solution of each metric under different values of ε_1 and ε_2. Fig. 7 shows the specific location coordinates of the optimal solution for each metric and representative calculation results with $\varepsilon_2 = 0.05$ and $\varepsilon_1 = 0.05 \sim 0.2$.

Combining the analysis of the results in Tables 2, 3 and Fig. 7, it can be seen that the calculation accuracy of RCTR and MRT has reached a very high standard. Due to CPD is a simulation metric, its calculation accuracy is lower than the former, but under different requirements of ε_1, the position of the optimal solution is also accurately found. Therefore, we believe that our simulation model and the DMLM are effective under different indicators and smaller values of ε_1 and ε_2.

5.2. Model sensitivity analysis

We use the topological and spatial data of the French optical fiber network to test the sensitivity of DMLM. The French backbone network has 88 nodes and 93 links, and its network deployment plane is set as a rectangle with a width of 580.83 km and length 1032.21 km. The network topology data comes from the website: the Internet Topology Zoo [47]. Values of equipment invulnerability (φ_v and φ_e) and repair conditions (μ_v, μ_e, μ_t, μ_l) are same with the above experiments. Considering the calculation speed and accuracy of the results, we set $\varepsilon_1 = 0.1$ and $\varepsilon_2 = 0.1$. The data source website did not give the specific type and bandwidth of the optical cable, and the official website of this

Table 2
Results of model validity analysis based on the linear function. The left column of each table cell is the average of the ratio of 100 calculation results to the optimal result, and the right column is the standard deviation.

Optimal solution	ε_1	ε_2							
		0.05		0.1		0.15		0.2	
RTCR γ^*=0.8682	**0.05**	0.9996	6.1e−05	0.9994	7.4e−05	0.9991	1.1e−05	0.9990	2.9e−05
	0.1	0.9994	1.6e−04	0.9992	1.9e−04	0.9990	1.3e−04	0.9992	5.3e−04
	0.15	0.9992	2.4e−04	0.9990	2.5e−04	0.9988	1.9e−04	0.9989	5.3e−04
	0.2	0.9990	3.3e−04	0.9988	2.9e−04	0.9987	2.8e−04	0.9986	5.0e−04
MRT T^*=8.4823	**0.05**	0.9992	8.5e−05	0.9952	5.1e−05	0.9913	3.5e−05	0.9881	2.1e−04
	0.1	0.9990	1.8e−04	0.9951	1.2e−04	0.9911	1.6e−04	0.9877	3.2e−04
	0.15	0.9988	3.3e−04	0.9949	2.5e−04	0.9909	3.0e−04	0.9874	4.86e−04
	0.2	0.9986	4.8e−04	0.9947	3.5e−04	0.9908	3.6e−04	0.9871	5.6e−04
CPD $Loss^*$=6.877	**0.05**	0.9049	1.5e−03	0.9120	1.8e−03	0.9191	1.9e−03	0.9248	2.2e−03
	0.1	0.9078	2.6e−03	0.9160	3.8e−03	0.9243	4.1e−03	0.9313	5.4e−03
	0.15	0.9094	2.7e−03	0.9185	4.3e−03	0.9277	5.8e−03	0.9359	7.7e−03
	0.2	0.9131	5.1e−03	0.9219	5.8e−03	0.9323	7.4e−03	0.9417	8.8e−03

Table 3
Results of model validity analysis based on the Gaussian distribution function. The left column of each table cell is the average of the ratio of 100 calculation results to the optimal result, and the right column is the standard deviation.

Optimal solution	ε_1	ε_2							
		0.05		0.1		0.15		0.2	
RTCR γ^*=0.8730	**0.05**	0.9994	1.0e−04	0.9984	1.0e−04	0.9991	8.3e−05	0.9980	8.3e−05
	0.1	0.9995	1.5e−04	0.9986	1.5e−04	0.9991	1.1e−04	0.9987	9.1e−04
	0.15	0.9995	1.5e−04	0.9986	1.9e−04	0.9992	1.7e−04	0.9990	1.1e−03
	0.2	0.9996	2.2e−04	0.9988	3.3e−04	0.9993	2.6e−04	0.9995	1.2e−03
MRT T^*=9.0377	**0.05**	0.9767	6.0e−04	0.9643	5.5e−04	0.9917	3.2e−02	0.9347	7.4e−04
	0.1	0.9761	1.1e−03	0.9637	8.9e−04	0.9662	3.7e−02	0.9341	1.1e−03
	0.15	0.9757	1.6e−03	0.9635	1.1e−03	0.9591	3.3e−02	0.9334	1.2e−03
	0.2	0.9750	1.8e−03	0.9630	1.5e−03	0.9550	3.1e−02	0.9332	1.6e−03
CPD $Loss^*$=8.0313	**0.05**	0.9179	1.3e−02	0.9343	1.4e−02	0.9443	1.9e−02	0.9625	2.0e−02
	0.1	0.9307	1.8e−02	0.9490	1.9e−02	0.9667	2.9e−02	0.9752	2.8e−02
	0.15	0.9362	2.2e−02	0.9708	3.1e−02	0.9913	3.5e−02	0.9989	4.1e−02
	0.2	0.9474	2.4e−02	0.9726	3.2e−02	0.9979	3.7e−02	0.9990	4.2e−02

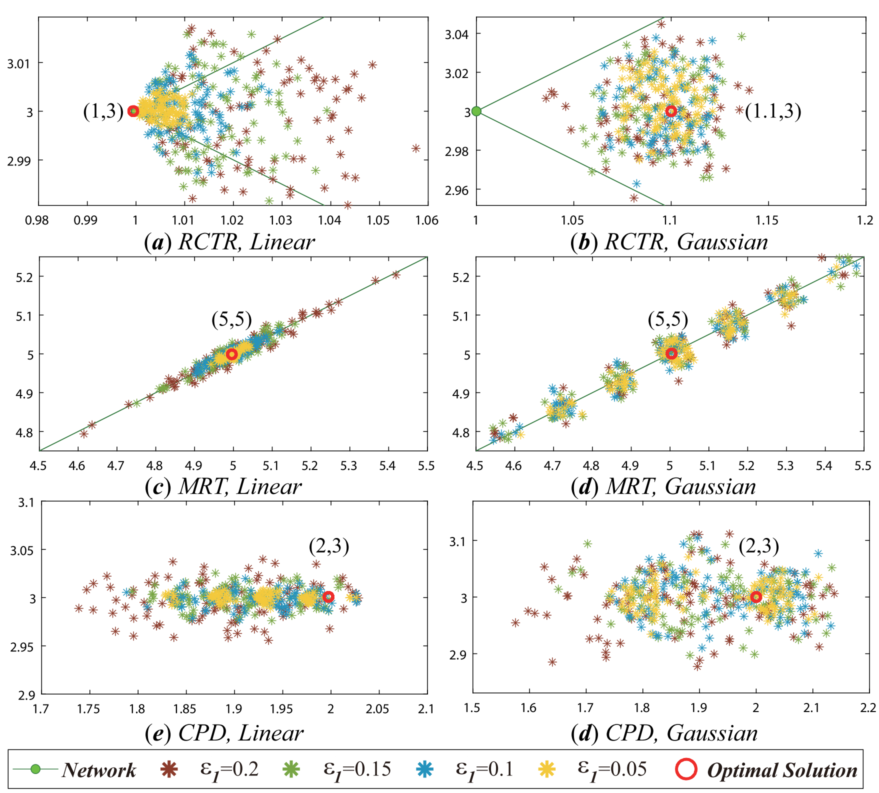

Fig. 7. Contrast between the locations of partial calculation results and locations of optimal solutions in model validity analysis. In all images, $\varepsilon_2 = 0.05$, and the calculation results of four different values of ε_1 are marked by four-color rice symbols. The red hollow circle is the position of the optimal solution. (a)(c)(e) show the results' positions of the three metrics under the linear function. (b)(d)(f) show the results' positions of the three metrics under the Gaussian distribution function.

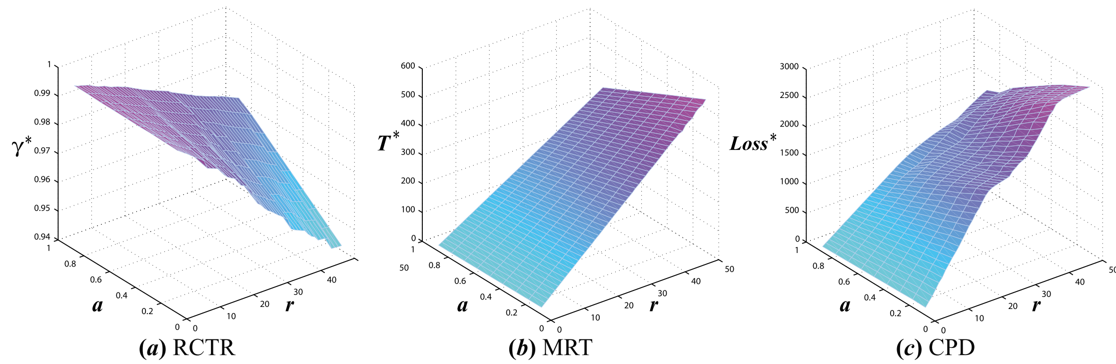

Fig. 8. Model sensitivity analysis results based on the linear function.(a) Optimal results of RCTR changes with damage parameters r and a. (b) Optimal results of MRT changes with damage parameters r and a. (c) Optimal results of CPD changes with damage parameters r and a.

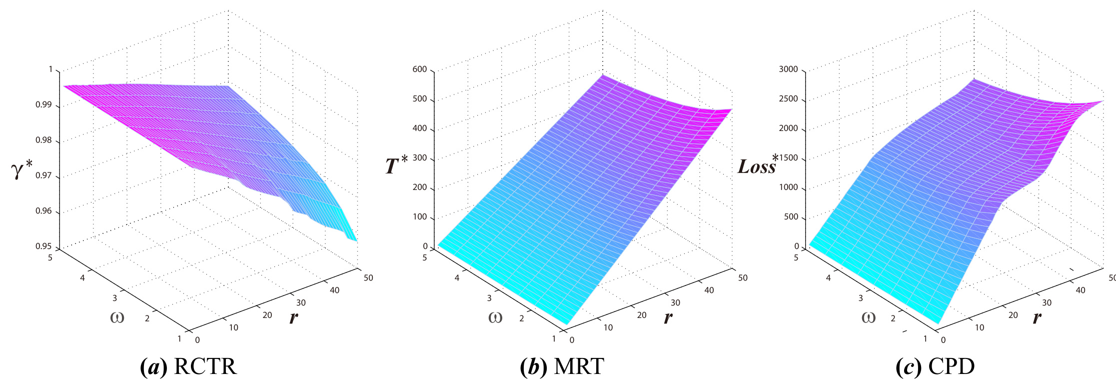

Fig. 9. Model sensitivity analysis results based on the Gaussian distribution function.(a) Optimal results of RCTR changes with damage parameters r and ω. (b) Optimal results of MRT changes with damage parameters r and ω. (c) Optimal results of CPD changes with damage parameters r and ω.

optical network operator did not give specific information. Based on the background that the production and operation time of the fiber network is relatively concentrated, we judge that the technical level of the equipment used in the network is not much different. According to the basic principles of fiber network design, we set $\forall F = \forall F_p = 1$ in an ideal situation.

This module is divided into two parts according to the two damage-efficiency functions. The first part tests the sensitivity of DMLM to important damage parameters r and a based on the linear function. We set the value range of the damage radius r to be $1 \sim 50$ km, and the damage efficiency a to be $0.1 \sim 0.9$, to test the changes of the extreme values of the three vulnerability metrics under actual fiber network background. The second part tests the sensitivity of DMLM to r and ω based on the Gaussian distribution function, and the damage efficiency ω to be $1 \sim 5$.

According to the Formula (12) and (37), the change trends of the optimal results of each metric in Fig. 8 with r and a are completely consistent with the formulas. The distributions of the results in Fig. 8(a)(b) prove that DMLM has excellent sensitivity under RCTR and MRT. Fig. 8(c) shows that the smoothness of CPD's result is visually worse than the others. It should be noted that there is still no singularity when $Loss^*$ increased with r and decreased with a. The reason for the jump of results may be that the network topology analysis is added in the CPD calculation process.

Fig. 9 shows the model sensitivity of the DMLM using the Gaussian distribution function. It can be seen that the DMLM still maintains a good sensitivity to the parameters ω and r. The value range of ω is much larger than a, and we only selected a small section for testing. Therefore, the change trend of the results shown in Fig. 9 is smaller than that in Fig. 8.

5.3. Comparison experiment analysis

The comparative experiments are the visual comparison of the vulnerability analysis results of RCTR with MRT and CPD, using the French optical fiber backbone network. We selected three specific damage scenes: The first scenario is $r = 5$ km $a = 0.4$ or $\omega = 1$, simulating the damage with small coverage and small attenuation(such as missile or bomb attacks). The second scenario is $r = 25$ km $a = 0.6$ or $\omega = 3$, simulating the damage with medium coverage and moderate damage attenuation (such as mass destruction weapons or hurricanes). The third scenario is $r = 50$ km $a = 0.8$ or $\omega = 5$, simulating the damage with large coverage and large attenuation (such as electromagnetic pulse bombs or earthquakes).

Fig. 10 illustrates the comparison of the visualization results of RCTR versus MRT and CPD in three damage scenarios based on the linear function. In Fig. 10, the double clustering of the candidate points in terms of geographic location and damage degrees appears, that is, some of the candidate points are located close to each other, and their damage degrees are similar. We define these clusters as vulnerable zones, and these zones are needed to be focused on network protection design and network daily maintenance. In the small damage scene, the vulnerable zones of RCTR are concentrated near the nodes with greater connection degree and are only distributed on the side close to links. In the large damage scene, the vulnerable area is more concentrated in geographic areas with densely distributed fiber links. In comparison, MRT's vulnerable zones are all distributed near optical fiber links. As the damage radius increases, MRT's vulnerable zones begin to concentrate on a small number of longer links, and the area of zones gradually increases. CPD is more like a combination of the two results: vulnerable zones are concentrated near optical fiber links that have the core position of the network topology. Compared with other metrics,

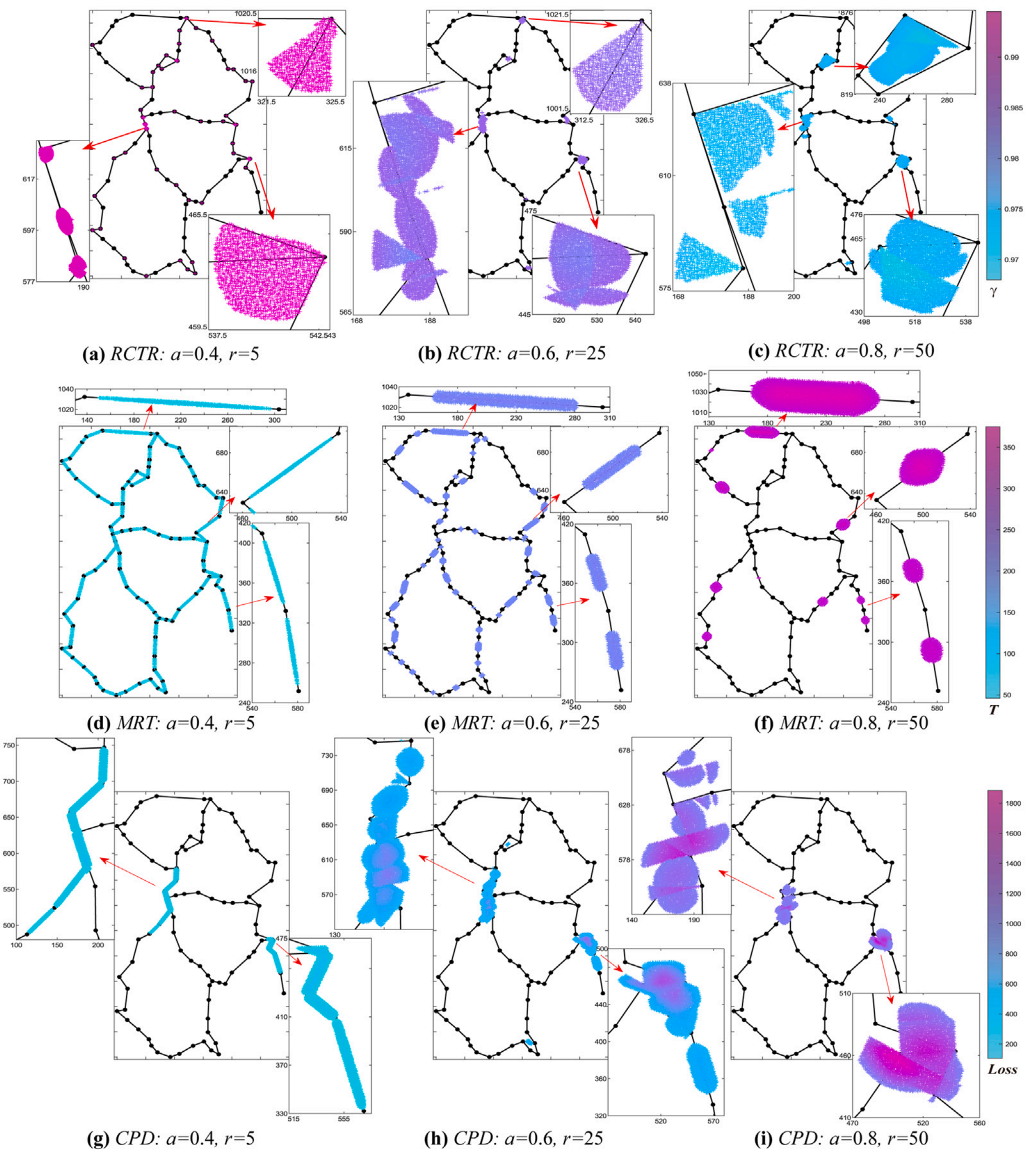

Fig. 10. Comparison of visualization results of RCTR vs MRT and CPD in three damage scenarios based on the linear function. (a)(b)(c) shows the locations and metric results of top 10000 candidate points under RCTR with three damage scenarios. (d)(e)(f) shows the locations and metric results of top 10000 candidate points under MRT with three damage scenarios. (f)(h)(i) shows the locations and metric results of top 10000 candidate points under CPD with three damage scenarios.

CPD can provide more accurate and effective recommendations for network protection design and daily maintenance under the small-scale damage situation.

Fig. 11 illustrates the comparison of visualization results based on the Gaussian distribution function. Comparing with Fig. 10, it can be seen that the vulnerable zones found by the Gaussian distribution function are very close to the linear function in terms of location and change trend, and differ in the areas and shapes. When CPD is used as the metric, the shape and area of vulnerable zones change the most before and after, which means that CPD is more sensitive to the damage-efficiency function. Therefore, it can be seen from the visual comparison charts of the vulnerable area, considering the network recoverability in the network vulnerability analysis would provide a new perspective for network design, protection, or maintenance.

6. Conclusion

Motivated by applications in optical fiber network survivability, we focused on the network vulnerability analysis. Considering the long repair time after fiber links are damaged, and the serious loss of network performance during the repair process, we believe that the traditional network vulnerability analysis method is no longer applicable to the fiber network. It is necessary to design a new simulation for the repair process of the damaged network, extract and calculate the mathematical characteristics displayed in the network repair process, to try to analyze the network vulnerability and find the vulnerable zone from a new perspective. Based on this motivation, we redesigned the fiber network simulation model and damage simulation model, which can accurately position the damage and its degree and is compatible with the repair simulation process. The proposal of the Link Simulation Model breaks the traditional link simulation method: two points determine one line, and re-simulates the link by taking virtual points at intervals, which solves the problem of precise positioning of the damaged position of fiber links. The proposal of the Damage-efficiency Circle Model solves the problem that the traditional probability damage model cannot describe the damage degree of network elements and provides a calculation basis for network repair simulation. We designed the Saturated Rescue Strategy for repair simulation. While maintaining the authenticity of the simulation, SRS does not require additional input

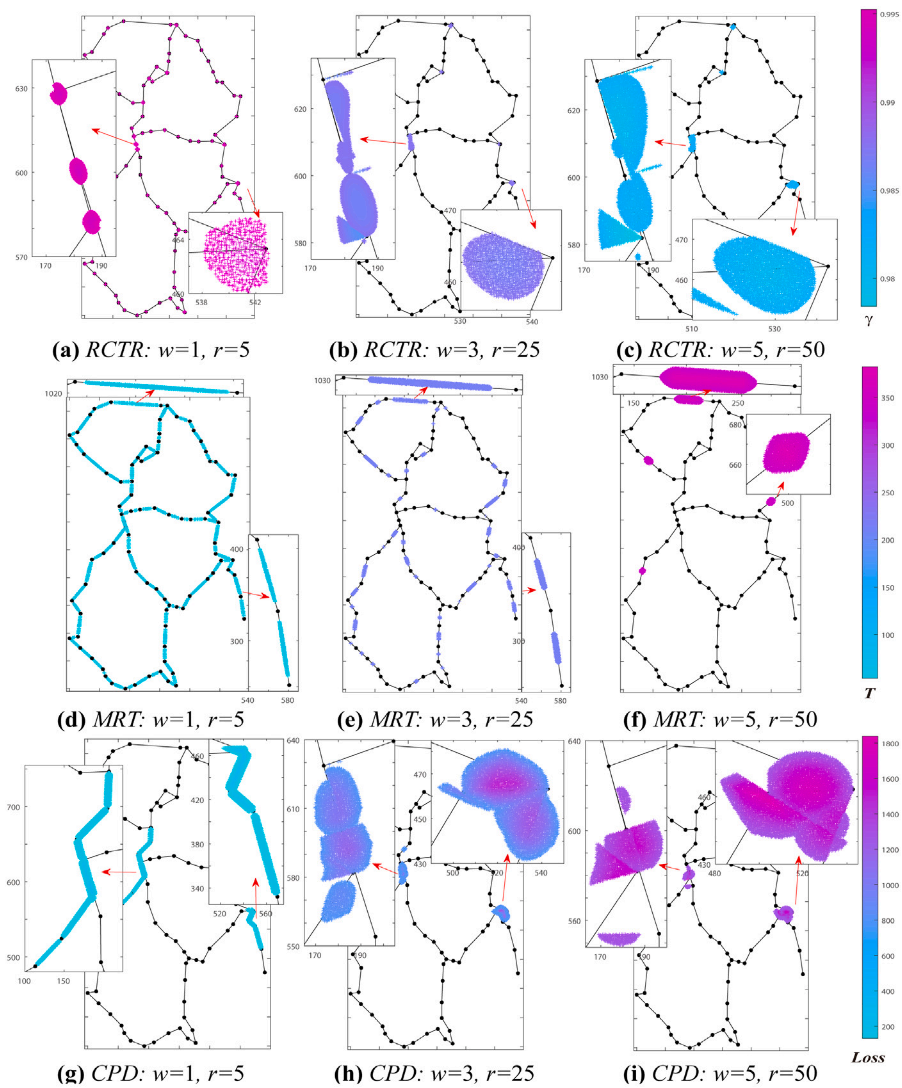

Fig. 11. Comparison of visualization results of RCTR vs MRT and CPD in three damage scenarios based on the Gaussian distribution function. (a)(b)(c) shows the locations and metric results of top 10000 candidate points under RCTR with three damage scenarios. (d)(e)(f) shows the locations and metric results of top 10000 candidate points under MRT with three damage scenarios. (f)(h)(i) shows the locations and metric results of top 10000 candidate points under CPD with three damage scenarios.

of data related to repair resources, which greatly improves the practicability and scope of application of the overall model. Based on the above simulation models, we designed the vulnerability analysis model: the Damage Measurement and Location Model. In DMLM, a heuristic traversal algorithm called Optimized Grid-partition-based method is proposed to locate the vulnerable zones of the fiber network with a faster calculation speed. We also designed supporting network performance metrics for DMLM. The related theoretical framework of each metric has also been developed to support the application of DMLM. Numerical results verify the effectiveness of the new model and prove that the model has excellent sensitivity to essential parameters. The visual results of vulnerable zones prove the importance of combining network recoverability with vulnerability analysis in network design, network protection, and network maintenance.

In the further study, we will further optimize the DMLM for the problem of how to combine the earthquake risk maps [48,49] with the vulnerability analysis of the optical fiber network. The dynamic repair model with time coordinates should be applied to network vulnerability analysis [50], to expand the application of different repair strategies in DMLM.

CRediT authorship contribution statement

Ke Wang: Conceptualization, Methodology, Software, Validation, Investigation, Formal analysis, Writing – original draft. **Jinfeng Liu:** Validation, Investigation, Resources. **Lai Tian:** Software, Validation. **Xianfeng Tan:** Resources, Writing – review & editing. **Guansheng Peng:** Writing – review & editing. **Tianwen Qin:** Resources, Supervision. **Jun Wu:** Conceptualization, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is partly supported by the National Natural Science Foundation of China under Grant No. 71871217 and the Natural Science Foundation of Hunan Province under Grant No. 2019JJ20019. The authors declared that they have no conflicts of interest to this work.

References

[1] Tanaka K, Yamazaki Y, Okazawa T, Suzuki T, Kishimoto T, Iwata K. Experiment on seismic disaster characteristics of underground cable. In: 14th world conference on earthquake engineering, 2008. 2008, https://www.researchgate.net/publication/267768959.

[2] Gomes T, Tapolcai J, Esposito C, Hutchison D, Kuipers F, Rak J, et al. A survey of strategies for communication networks to protect against large-scale natural disasters. In: 2016 8th international workshop on resilient networks design and modeling. IEEE; 2016, p. 11–22.

[3] Carter L, Gavey R, Talling PJ, Liu JT. Insights into submarine geohazards from breaks in subsea telecommunication cables. Oceanography 2014;27(2):58–67.

[4] Wilson C. High altitude electromagnetic pulse (HEMP) and high power microwave (HPM) devices: Threat assessments. In: Library of congress Washington DC congressional research service, 2007. 2007, http://www.todaysengineer.org/2007/Sep/HEMP.asp.

[5] Wang XL, Jiang XH, Pattavina A. Assessing network vulnerability under probabilistic region failure model. In: 2011 IEEE 12th international conference on high performance switching and routing. IEEE; 2011, p. 164–70.

[6] Kobayashi M. Experience of infrastructure damage caused by the great east Japan earthquake and countermeasures against future disasters. IEEE Commun Mag 2014;52(3):23–9.

[7] Habib MF, Tornatore M, Dikbiyik F, Mukherjee B. Disaster survivability in optical communication networks. Comput Commun 2013;36(6):630–44.

[8] Lü LY, Chen DB, Ren XL, Zhang QM, Zhang YC, Zhou T. Vital nodes identification in complex networks. Phys Rep 2016;650:1–63.

[9] Galvan G, Agarwal J. Assessing the vulnerability of infrastructure networks based on distribution measures. Reliab Eng Syst Saf 2020;196:106743.

[10] Xiong W, Zhou J, Zhuo M, Li X, Xu J. Military grand cihai part 2. Great Wall Publishing House; 2000.

[11] Agarwal PK, Efrat A, Ganjugunte SK, Hay D, Sankararaman S, Zussman G. The resilience of WDM networks to probabilistic geographical failures. IEEE/ACM Trans Netw 2013;21(5):1525–38.

[12] Dikbiyik F, Tornatore M, Mukherjee B. Minimizing the risk from disaster failures in optical backbone networks. J Lightwave Technol 2014;32(18):3175–83.

[13] Ran Y. Considerations and suggestions on improvement of communication network disaster countermeasures after the wenchuan earthquake. IEEE Commun Mag 2011;49(1):44–7.

[14] Morrison KT. Rapidly recovering from the catastrophic loss of a major telecommunications office. IEEE Commun Mag 2011;49(1):28–35.

[15] Al Sabeh K, Tornatore M, Dikbiyik F. Progressive network recovery in optical core networks. In: 2015 7th international workshop on reliable networks design and modeling. IEEE; 2015, p. 106–11.

[16] Pourvali M, Liang KL, Feng G, Hao B, Shaban K, Khan S, et al. Progressive recovery for network virtualization after large-scale disasters. In: International conference on computing. IEEE; 2016, p. 1–5.

[17] Genda K. Effective network recovery with higher priority to service recovery after a large-scale failure. In: 2017 23rd Asia-pacific conference on communications. IEEE; 2017, p. 1–6.

[18] Tootaghaj DZ, La Porta T, He T. Modeling, monitoring and scheduling techniques for network recovery from massive failures. In: 2019 IFIP/IEEE symposium on integrated network and service management. IM, IEEE; 2019, p. 695–700.

[19] Liem AT, Hwang IS, Nikoukar AA. An autonomous recovery mechanism against optical distribution network failures in EPON. Opt Fiber Technol, Mater Devices Syst 2014;20(5):552–60.

[20] Zhang X, Guo L, Hou WG, Zhang QH, Wang SQ. Failure recovery solutions using cognitive mechanisms based on software-defined optical network platform. Opt Eng 2017;56(1):016107.

[21] Dikbiyik F. Multi-repairmen problem for disaster recovery of optical networks. Sakarya ÜNiversitesi Fen Bilimleri EnstitÜSÜ Dergisi 2017;21(1):47–53.

[22] Ma C, Colman-Meixner C, Tornatore M, Zhao YL, Zhang J, Mukherjee B. Multiple traveling repairmen problem with virtual networks for post-disaster resilience. In: 2016 IEEE international conference on communications. IEEE; 2016, p. 1–6.

[23] Hosseini S, Barker K, Ramirez-Marquez JE. A review of definitions and measures of system resilience. Reliab Eng Syst Saf 2016;145:47–61.

[24] Cao QD, Miles SB, Choe Y. Infrastructure recovery curve estimation using Gaussian process regression on expert elicited data. Reliab Eng Syst Saf 2022;217:108054.

[25] Saadat Y, Ayyub BM, Zhang Y, Zhang D, Huang H. Resilience of metrorail networks: Quantification with washington, DC as a case study. ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg 2019;5(4):1–14.

[26] Shafieezadeh A, Burden LI. Scenario-based resilience assessment framework for critical infrastructure systems: Case study for seismic resilience of seaports. Reliab Eng Syst Saf 2014;132:207–19.

[27] Ouyang M, Liu C, Xu M. Value of resilience-based solutions on critical infrastructure protection: Comparing with robustness-based solutions. Reliab Eng Syst Saf 2019;190:106506.

[28] Mottahedi A, Sereshki F, Ataei M, Qarahasanlou AN, Barabadi A. Resilience estimation of critical infrastructure systems: Application of expert judgment. Reliab Eng Syst Saf 2021;215:107849.

[29] Cheng Y, Elsayed E, Chen X. Random multi hazard resilience modeling of engineered systems and critical infrastructure. Reliab Eng Syst Saf 2021;209:107453.

[30] Arvidsson B, Johansson J, Guldåker N. Critical infrastructure, geographical information science and risk governance: A systematic cross-field review. Reliab Eng Syst Saf 2021;213:107741.

[31] Phillips CA. The network inhibition problem. In: Proceedings of the twenty-fifth annual ACM symposium on theory of computing. DBLP; 1993, p. 776–85.

[32] Albert R, Jeong H, Barabási A-L. Error and attack tolerance of complex networks. Nature 2000;406(6794):378–82.

[33] Neumayer S, Zussman G, Cohen R, Modiano E. Assessing the vulnerability of the fiber infrastructure to disasters. IEEE/ACM Trans Netw 2011;19(6):1610–23.

[34] Agarwal PK, Efrat A, Ganjugunte SK, Hay D, Sankararaman S, Zussman G. Network vulnerability to single, multiple, and probabilistic physical attacks. In: 2010 military communications conference. IEEE; 2010, p. 1824–9.

[35] Xuan Y, Shen YL, Nguyen NP, Thai MT. Efficient multi-link failure localization schemes in all-optical networks. IEEE Trans Commun 2013;61(3):1144–51.

[36] Banerjee S, Shirazipourazad S, Ghosh P, Sen A. Beyond connectivity-new metrics to evaluate robustness of networks. In: 2011 IEEE 12th international conference on high performance switching and routing. IEEE; 2011, p. 171–7.

[37] Long XL, Tipper D, Gomes T. Measuring the survivability of networks to geographic correlated failures. Opt Switch Netw 2014;14:117–33.

[38] Wang K, Li Y, Wu J. Identifying multiple vulnerable areas of infrastructure network under global connectivity measure. Internat J Modern Phys C 2019;30(7):241–4.

[39] Neumayer S, Efrat A, Modiano E. Geographic max-flow and min-cut under a circular disk failure model. Comput Netw 2015;77:117–27.

[40] Trajanovski S, Kuipers FA, Mieghem PV, Ili A, Crowcroft J. Critical regions and region-disjoint paths in a network. In: Ifip networking conference. IEEE; 2013, p. 1–9.

[41] Barthelemy M. Spatial network. Phys Rep 2010;499(1).

[42] Xiao Z, Wang XL, Jiang XH, Lu SL. Degree of network damage: A measurement for intensity of network damage. In: 2014 19th European conference on networks and optical communications. IEEE; 2014, p. 140–6.

[43] Fan SL, Zhang YF, Yao T, Min ZX. Research summary of battlefield damage assessment methods on arms and equipments. J Acad Armored Force Eng 2013;27(1):21–6.

[44] Ma C, Zhang J, Zhao YL, Habib MF, Savas SS, Mukherjee B. Traveling repairman problem for optical network recovery to restore virtual networks after a disaster. IEEE/OSA J Opt Commun Networking 2015;7(11):81–92.

[45] Wu J, Barahona M, Tan YJ, Deng HZ. Spectral measure of structural robustness in complex networks. IEEE Trans Syst Man Cybern A: Syst Hum 2011;41(6):1244–52.

[46] De Iuliis M, Kammouh O, Cimellaro GP, Tesfamariam S. Quantifying restoration time of power and telecommunication lifelines after earthquakes using Bayesian belief network model. Reliab Eng Syst Saf 2021;208:107320.

[47] Knight S, Bowden R, Roughan M, Nguyen HX, Falkner N. The internet topology zoo. IEEE J Sel Areas Commun 2011;29(9):1765–75.

[48] Agrawal A, Bhatia V, Prakash S. Network and risk modeling for disaster survivability analysis of backbone optical communication networks. J Lightwave Technol 2019;37(10):2352–62.

[49] Chen W, Zhang L. Resilience assessment of regional areas against earthquakes using multi-source information fusion. Reliab Eng Syst Saf 2021;215:107833.

[50] Yang L, Ye ZS, Chi-Guhn L, Yang SF, Peng R. A two-phase preventive maintenance policy considering imperfect repair and postponed replacement. European J Oper Res 2019;274(3):966–77.