Disaster resilience of optical networks: State of the art, challenges, and opportunities

Jacek Rak a,*, Rita Girão-Silva b, c, Teresa Gomes b, c, Georgios Ellinas d, Burak Kantarci e, Massimo Tornatore f

a Gdańsk University of Technology, Faculty of Electronics, Telecommunications and Informatics, G. Narutowicza 11/12, 80-233, Gdańsk, Poland
b University of Coimbra, Department of Electrical and Computer Engineering, 3030-290, Coimbra, Portugal
c Instituto for Systems Engineering and Computers at Coimbra (INESC Coimbra), 3030-290, Coimbra, Portugal
d University of Cyprus, Department of Electrical and Computer Engineering, KIOS Research and Innovation Center of Excellence, School of Engineering, 75 Kallipoleos Street, P.O. Box 20537, 1678, Nicosia, Cyprus
e University of Ottawa, School of Electrical Engineering and Computer Science, 800 King Edward Avenue, Ottawa, ON, K1N 6N5, Canada
f Politecnico di Milano, Department of Electronics, Information and Bioengineering, Via Pontisio 34, 20133, Milano, Italy

ARTICLE INFO

Keywords:
Optical networks
Resilience
5G and beyond
Content connectivity
Data evacuation
Datacenters
Failure recovery
Fixed-mobile convergence
Malicious attacks
Manycast
Natural disasters
Network availability
Post-disaster modeling
Post-disaster recovery
Power disruptions
Weather-based disruptions

ABSTRACT

For several decades, optical networks, due to their high capacity and long-distance transmission range, have been used as the major communication technology to serve network traffic, especially in the core and metro segments of communication networks. Unfortunately, our society has often experienced how the correct functioning of these critical infrastructures can be substantially hindered by massive failures triggered by natural disasters, weather-related disruptions and malicious human activities.

In this position paper, we discuss the impact on optical networks of all major classes of disaster events mentioned above, and we overview recent relevant techniques that have been proposed to increase the disaster resilience of optical networks against the various classes of disaster events. We start by presenting some proactive methods to be applied before the occurrence of a disaster. Then we move our focus also on other preparedness methods that can be executed in the (typically short) time frame between the occurrence of an early alert of an incoming disaster and the time a disaster actually hits the network. Finally, we discuss reactive procedures that allow performing post-disaster recovery operations effectively. The analysis of disaster resilience mechanisms provided in this paper covers both wired and optical wireless communication infrastructures and also contains explicit remarks covering the role of emerging technologies (e.g., fixed-mobile convergence in the 5G era and beyond) in disaster resilience.

1. Introduction

Communication networks are critical infrastructures that provide end users with several fundamental network services in today’s society, as remote working, e-banking and e-government services, or smart grid management. Therefore, providing reliable communication services is of the utmost importance to guarantee adequate access to these fundamental services. For several decades, optical networks have represented the major communication infrastructure thanks to their high capacity and the long-distance transmission range. However, their correct functioning is threatened by a diverse set of possible disaster events that can lead to concurrent failures of multiple network elements. As shown in Fig. 1, the major external causes of massive failures in optical networks can be categorized as natural disasters, weather-related events and malicious human activities.

Natural disasters include predictable events (e.g., floods, fires, hurricanes, tornadoes, volcano eruptions) and unpredictable events such as earthquakes [1] – all leading to permanent failures of network nodes/links. For instance, the 7.1-magnitude earthquake in December 2006 in Taiwan disrupted seven submarine optical links and significantly degraded Internet connectivity between Asia and North America for weeks [2,3]. Similarly, the 2005 hurricane Katrina that hit southern US resulted in the failure of

* Corresponding author.

E-mail addresses: jrak@pg.edu.pl (J. Rak), rita@deec.uc.pt (R. Girão-Silva), teresa@deec.uc.pt (T. Gomes), gellinas@ucy.ac.cy (G. Ellinas), Burak.Kantarci@uottawa.ca (B. Kantarci), massimo.tornatore@polimi.it (M. Tornatore).

https://doi.org/10.1016/j.osn.2021.100619
Received 20 July 2020; Accepted 2 March 2021
Available online 26 March 2021
1573-4277/© 2021 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license.
several major network nodes that lasted over ten days due to concurrent power grid outages [4]. It is worth noting that, as disasters commonly hit specific and geographically-circumscribed areas, disaster failures in optical networks are also commonly referred to as regional failures [5].

As optical communication systems are currently being successfully applied also in the wireless domain (optical wireless communications – OWC), e.g., as a backup architecture for optical wired networks or in areas where OWC is the only choice (consider inter-satellite links, space to ground communications, or communications in dense rural areas), the transient effects of weather-induced disruptions are addressed as well in this paper. In particular, dense fog, snow, and clouds are major factors leading to partial (or a complete) degradation of OWC links capacity [6,7]. Although such events do not result in physical failures of network elements and the networks recover automatically once these weather factors subside, due to the frequency and duration of adverse weather conditions and their intensification in certain areas, there is a need to apply proper mechanisms to guarantee network resilience under such circumstances.

Massive failures in optical networks can also be the result of malicious human activities. In these cases, the network is the primary target of an attacker aiming to cause the most severe damage possible. Some examples of malicious human activities include, electromagnetic pulse (EMP) attacks [8], distributed denial of service (DDoS) attacks [9], or the indirect effect of an attack performed through weapons of mass destruction and affecting the optical network.

As disaster events are currently growing in number, intensity and scale [10,11], disaster resilience is also a topic of growing importance. In particular, as failures in optical networks cannot be eliminated entirely [7], and since the standard protection mechanisms – adequate for assuring the resilience under failures of single nodes and links – are not suitable in case of large-scale disasters [12], it is of the utmost importance to investigate and deploy mechanisms focused on network disaster resilience. In Ref. [7], disaster resilience is defined as the ability of the network to provide and maintain an acceptable level of service in the face of disaster-induced faults and challenges to normal operation.

As the consequences of disasters are often severe, and are often intensified by the fact that people increase their network activities in face of disaster-induced faults and challenges to normal operation.

As disaster events are currently growing in number, intensity and scale [10,11], disaster resilience is also a topic of growing importance. In particular, as failures in optical networks cannot be eliminated entirely [7], and since the standard protection mechanisms – adequate for assuring the resilience under failures of single nodes and links – are not suitable in case of large-scale disasters [12], it is of the utmost importance to investigate and deploy mechanisms focused on network disaster resilience. In Ref. [7], disaster resilience is defined as the ability of the network to provide and maintain an acceptable level of service in the face of disaster-induced faults and challenges to normal operation.

As the consequences of disasters are often severe, and are often intensified by the fact that people increase their network activities in face of disaster-induced faults and challenges to normal operation.

As disaster events are currently growing in number, intensity and scale [10,11], disaster resilience is also a topic of growing importance. In particular, as failures in optical networks cannot be eliminated entirely [7], and since the standard protection mechanisms – adequate for assuring the resilience under failures of single nodes and links – are not suitable in case of large-scale disasters [12], it is of the utmost importance to investigate and deploy mechanisms focused on network disaster resilience. In Ref. [7], disaster resilience is defined as the ability of the network to provide and maintain an acceptable level of service in the face of disaster-induced faults and challenges to normal operation.

As the consequences of disasters are often severe, and are often intensified by the fact that people increase their network activities in face of disaster-induced faults and challenges to normal operation.

As disaster events are currently growing in number, intensity and scale [10,11], disaster resilience is also a topic of growing importance. In particular, as failures in optical networks cannot be eliminated entirely [7], and since the standard protection mechanisms – adequate for assuring the resilience under failures of single nodes and links – are not suitable in case of large-scale disasters [12], it is of the utmost importance to investigate and deploy mechanisms focused on network disaster resilience. In Ref. [7], disaster resilience is defined as the ability of the network to provide and maintain an acceptable level of service in the face of disaster-induced faults and challenges to normal operation.

(i) before a disaster, by using proactive mechanisms, e.g., pre-allocation of the alternative transmission paths in case of failure occurrence,

(ii) in the (typically short) time frame between the reception of a disaster alert and the actual occurrence of a disaster (for instance by evacuating critical data/service just before an incoming hurricane actually reaches the region where the network infrastructure is located),

(iii) after a disaster, by using reactive techniques for post-disaster recovery that can maximize network utility in the aftermath of a disaster, e.g., by installing temporary wireless links to replace the affected wired links.

In the rest of the paper, relevant mechanisms for disaster resilience in optical networks which can be applied in each of these phases will be discussed.

1.1. Goal and organization of the paper

The main goal of this paper is to present the most important mechanisms to assure disaster resilience in optical networks, including:

(a) proactive protection mechanisms for optical wired networks against failures induced by unpredictable disasters (mainly, earthquakes),

(b) preparedness for foreseeable disasters by mitigation strategies used just after the disaster symptoms occurrence,

(c) resilience of optical wireless networks under weather-induced disruptions,

(d) mitigation of the impact of malicious attacks,

(e) post-disaster recovery of the optical communication architecture and its services.

This paper is intended to provide an extensive survey of the topic, but it targets an up-to-date and a comprehensive view on the latest advances on the important topic of disaster resilience of optical networks. Indeed, despite the availability of several survey papers addressing specific problems (see for example [15,16]), in our opinion the set of techniques presented in those papers is limited. In particular, none of them addresses a complete range of recent up-to-date major problems (a)–(e) as highlighted above, together with the respective solutions.

For instance, in addition to Ref. [16], our paper presents protection and restoration mechanisms classified according to the type of a natural disaster (in particular, whether a disaster is foreseeable or not), weather-induced disruptions or human-induced attacks, as well as it provides description of selected major pre-planned data evacuation strategies along with characteristics of the post-disaster recovery schemes.

Ref. [15], which provides a comprehensive set of disaster-related problems for optical networks, was published in 2013, and it currently misses the description of all the relevant schemes proposed after its publication date in the last 8 years.

The scope of the survey paper [17] is limited to disaster recovery techniques based on the Virtual Machine (VM) live migration and the related problems of transferring large amount of data over long distances among datacenters (DCs), to increase the number of protected resources. As highlighted above, the scope of our paper is broader, and this problem is only one of issues addressed by our paper.

The paper [10] focuses on the presentation of vulnerabilities of communication networks to disaster-induced disruptions. In this context, it presents a selected set of measures of network vulnerability, the overview of methods of identification of vulnerable regions, and the overview of vulnerability of physical infrastructures. The latter part of that paper highlights a selected set of pre-disaster and post-disaster recovery techniques, as well as provides the overview of selected

![Fig. 1. Classification of major external events responsible for massive failures in optical networks.](image-url)
algorithms of disaster-resilient routing. Compared to Ref. [10], our paper provides a more detailed and up-to-date analysis in each of the aspects mentioned above with a particular focus on optical communications, as well as extends the scope of analysis, e.g., by a section on optical wireless architectures.

The remaining part of this paper is organized as follows. Section 2 describes the general assumptions and requirements for disaster resilience in optical networks covering all five major contexts (a)–(e) mentioned above. Next, in Section 3, we present the most important concepts related to disaster resilience of optical networks to earthquakes (issue (a) above) being the major type of unpredictable natural disasters affecting the optical wired-wide-area networks. Section 4 highlights the concept of preparedness for foreseeable disasters – issue (b), for which the respective recovery operations (such as data evacuation) can be performed just-in-time in the period after identifying the symptoms of a disaster and before it hits the network. Resilience of optical wireless communications under adverse weather conditions (issue (c) above) is, in turn, addressed in Section 5. Section 6 focuses on the presentation of techniques for mitigating the impact of malicious human activities (d) on the performance of optical communication systems, while Section 7 addresses the problems and selected solutions allowing the post-disaster recovery of the optical communication architecture and its services (issue (e) above). Section 8 concludes the paper.

2. Related works

Resilience of communication networks has often been the topic of discussions in the related literature. For instance, frameworks for network resilience are addressed in Ref. [11]. In general, the definition of resilience, presented in the previous section of this paper in the context of disaster-induced massive failures, originally subsumes multiple related disciplines categorized into challenge tolerance and trustworthiness disciplines (see the work [14] by Sterbenz et al.). Challenge tolerance (comprising survivability, traffic tolerance and disruption tolerance) focuses on the design and engineering of resilient networks. Trustworthiness, in turn, includes dependability, security and performability, and is aimed at describing the resilience of networked systems [7,14].

A survey on resilience strategies concerning the large-scale natural disasters is presented in Ref. [10]. In Ref. [18], the authors tackle the problem of weather-based disruptions both for wireless networks (which are mainly affected by rain, snow, fog or wind) and for wired networks (which are mainly affected by foreseeable disasters such as floods caused by heavy rain).

In Ref. [15], a critical review of works on modeling and surviving multiple correlated disaster failures in optical Wavelength Division Multiplexing (WDM) mesh networks is put forward. Disasters are classified based on their features and their impact on communication networks. Strategies to mitigate that impact are reviewed. In Ref. [1], the focus is on the provisioning of the content to all the elements in a network in the event of disasters. As networks are becoming content-centric, the emphasis should be on guaranteeing that even if a network becomes partitioned, all the network portions have managed to replicate the necessary contents. This way, if the characteristics of disasters and consequential failures are taken into account, the networks may be prepared in advance to assure provisioning of services and a quick recovery of the network. In Ref. [16], several studies on issues associated with protection, restoration, cascading failures, disaster-based failures, and congestion-aware routing in optical networks, are reviewed. In particular, the problems of simultaneous cascading failures and disaster-aware network survivability are described. Some challenges regarding disaster-resilient network survivability are put forward.

Following [19], multidisciplinary approaches are often necessary to deal with the problems associated with disasters. In a disaster scenario (either natural or human-caused), communications acquire an even greater importance due to its necessity not only for the citizens but also for the relief teams. As service disruptions caused by disasters may be severe, according to Ref. [20] they raise the need to:

- (i) identify the vulnerabilities and prepare the networks to be able to react to possible disasters when those disasters occur – referred to as the disaster preparedness;
- (ii) have a framework in place that will allow for a rapid repair of the network and a quick provision of services – initial post-disaster response;
- (iii) establish rehabilitation and recovery plans.

In the remaining part of this section, we address the impact of failures of telecommunication infrastructures, and summarize the recommendations and best-practice guidelines of the Telecommunication Standardization Sector of The International Telecommunications Union (ITU-T) for disaster mitigation and relief. Then we discuss the geographically correlated failures and present the relevant approaches which allow to enhance the resilience of optical networks against this type of disaster failures. After that, the related risk analysis and models for natural disasters are described. This section is concluded with a brief overview of malicious attacks occurring in optical networks and of scenarios justifying the need to apply additional disaster resilience mechanisms in the post-disaster period.

2.1. Failures in communication infrastructures

The assessment of resilience of a certain infrastructure in the event of disasters may be quantitatively represented based on quality functions, which are used to model the recovery of a system in time [21]. In Ref. [21], the operability of a system in time t is modeled by a function involving the post-event operability, the operability of the system in regular conditions and a parameter related to the convergence of the system to its functional state. The value of these parameters is based on observations and information gathered in the aftermath of a disaster, during the disaster response and the recovery stages.

The ITU-T has issued a number of reports on the use, resilience and recovery of public telecommunication network infrastructures during disasters [20,22]. A report on the standardization efforts of the ITU-T for resilient Information and Communication Technologies (ICTs) is given in Ref. [23]. The impact of different disasters on the public telecommunication systems is described along with a revision of some recommendations in the topic of emergency telecommunications. The report [24] lists several generic recommendations and best-practice guidelines in terms of telecommunications for disaster mitigation and relief. As illustrated in Fig. 2, the most important recommendations include:

- in the preparedness stage: (i) it is advisable to have disaster plans updated according to lessons learned from simulations and past real disasters; (ii) regular drills to test the plans must be undertaken, which should help in their improvement; (iii) pre-positioning of the equipment and mobile resource units across hazardous areas; (iv) appropriate training of personnel and citizens.

In the particular case of telecommunications, it is important to have resilient and redundant equipment so that the alternative solutions may be quickly implemented. The appropriate location of the restoration equipment and of the emergency response facilities may be formulated as a stochastic programming model [25]. The diversity of means of connectivity is also important, as different networks (e.g., satellite, radio services) may be affected by disasters. One of the possible strategies for setting up the resilient ICT architecture in the aftermath of a disaster is the use of a Movable and Deployable ICT Resource Unit (MDRU) [26,27], which must be part of the disaster preparedness efforts. An MDRU is a portable unit with the necessary equipment to quickly deploy a local Wi-Fi network for provisioning the minimum necessary ICT services in disaster areas.
This may include using any remaining optical fiber cables and/or satellite communication lines to connect to other functional networks.

- in the immediate aftermath of the disaster: (i) policies regarding simplification of procedures (for instance licensing procedures for equipment or visa requirements for aid personnel) may help in assembling teams to deal with the immediate needs of citizens; (ii) making sure the relief teams are given appropriate and up-to-date information, which requires immediate steps to establish some form of communications; (iii) fostering collaborations between the different players to facilitate the rescue efforts.

In the particular case of telecommunications, it is important to start by assuring communications for the relief teams, but also for the general population. The other supporting infrastructures (e.g., power networks) also need to become operational in a short time frame.

- in the rehabilitation and recovery stage: (i) incorporate emerging technologies that may help in the future emergency preparedness and response planning; (ii) keep improving knowledge and learning from other disastrous events.

Concerning other recommendations, report [28] focuses on urban disasters and presents a comprehensive analysis of the possible failures that may occur in telecommunication infrastructures, namely the physical destruction of network components, the disruption in other infrastructures that support telecommunications (e.g., power networks), and the service disruptions due to network congestion or overload. These possible causes for failures are also analyzed in Ref. [29]. Another aspect investigated in Ref. [28] refers to the consequences of telecommunication failures in four different stages of post-disaster recovery (during the emergency response, the initial repairs, the reconstruction, and finally during the redevelopment of the affected structures). The authors also mention the selected aspects for strengthening and preparing the infrastructures for future disasters.

2.2. Geographically correlated failures

The link and node failures caused by disasters usually have a geographical correlation. Several metrics, proposed in the literature to quantify the impact of geographically correlated failures, are briefly reviewed in Ref. [30], and, in particular, the maximum variation of the weighted spectrum of a graph is proposed to assess the survivability of the network to geographically correlated failures. This new measure was shown to be quite versatile, as it could be used in unweighted and weighted graphs to evaluate geographically correlated vulnerable links and vulnerable nodes.

A stochastic model of geographically correlated link failures caused by disasters is proposed in Ref. [31], from which the probabilistic information regarding the failure of a set of links or the disconnection of nodes may be devised. With this information, an analysis of the availability of different services in the event of disaster failures may be performed. In the situation, when it is not possible to pre-determine the effect of disasters, a probabilistic study needs to be undertaken [32]. Scenarios of simultaneous attacks or attacks on interdependent components help in identifying the vulnerable points in the network, i.e., elements requiring the additional protection.

The network resilience to disasters entailing a geographical impact may be accomplished by considering a pair of geodiverse primary and backup paths for end-to-end connections. This way, whenever a disaster strikes and affects one or more intermediate elements of the primary path, the backup path (that should be sufficiently distant from the primary path) probably has not been affected and may be used. Several metrics may be used to assess the geodiversity, i.e., the geographical distance of two paths. For instance, in Ref. [33], the distance between two paths is the minimal distance between any nodes in those paths (except the source and the destination nodes), as presented in Fig. 3 (inspired by [33, Fig. 2]). Another metric is considered in Ref. [34] involving the minimal distance between an intermediate element (node or link) in one path and an intermediate element in the other path.

In Ref. [35] the problem of path geodiversity is tackled alongside the problem of end-to-end availability, i.e., the required levels of availability and geodiversity must be attained, and for that the availability of some edges may have to be increased. The edges with upgraded availability are selected in order to minimize the upgrade cost, which is important for network operators. Two different heuristics (one based on an Integer Linear Programming (ILP) approach and the other based on a greedy approach) are described and computational results are presented.

The concept of path diversification established in Ref. [36] is the basis for the definition of geodiverse paths, later considered, e.g., in Refs. [33,37]. In Ref. [36], the authors select the paths based on a diversity measure that allows for low latency and maximum flow reliability. In Ref. [33], the vulnerability of networks to cascading and regional-correlated failures due to natural disasters and intentional attacks is explored. A graph resilience metric, compensated Total
Geographical Graph Diversity (cTGGD), is proposed to provide information on the resilience level of different optical fiber networks. A routing protocol, GeoDivRP, is presented with two heuristics to calculate the geographically separated paths. The performance of GeoDivRP is improved in Ref. [37] by including information on critical regions identified with the proposed model.

A Shared Risk Link Group (SRLG) may be considered as a model for disasters affecting a few elements in a certain region of a network. It is defined as a set of links with a common risk of failure (e.g., fibers sharing a cable or a duct) [38]. If two paths are SRLG-disjoint, they do not share a common risk of failure. In Ref. [39], maximally SRLG-disjoint geodiverse paths are considered, again to ensure resilience to geographically-correlated disasters. Comparing to a situation when only link-disjointness is considered, the obtained path lengths tend to be higher and the cost of the required transponders is also higher, but this is necessary to provide the appropriate resilience. Specific strategies for the reduction of the cost of transponders were put forward, so that the cost for network operators does not become excessive. The observed results show that SRLG-disjointness already assures some degree of geodiversity for most end-to-end connections.

In Ref. [40], a FRAmework for DIsaster Resilience (FRADIR) is proposed, where network design, failure modeling and protection routing are investigated to improve the disaster resilience of mission-critical applications. In the design of the network, some links forming a spanning tree are selected for an upgrade of their availability, so as to achieve a higher availability substructure. The modeling of regional failures is accomplished by using SRLGs, in which the failure probability of a component depends both on the geographical distance from a disaster area and on the component availability. Finally, the paths for routing are selected using the General Dedicated Protection (GDP) method [41]. An extension of FRADIR is put forward in Ref. [42], where a heuristic algorithm is used to further upgrade the availability of the links to minimize the probability of failures disconnecting the network. A further evolution of FRADIR is proposed in Ref. [43], namely in terms of the steady-state network planning (a two-stage resolution approach was used to achieve a desired target availability threshold at minimal cost) and a more realistic disaster model for earthquakes (considered in failure modeling).

2.3. Risk analysis and models for natural hazards

Researchers in differentiated areas (e.g., climatology, geology) have developed models for natural hazards, such as wildfires, different geological events (including earthquakes and volcanic eruptions), different atmospheric events (including strong wind, heavy rain, ice and snow) and floods. Modeling natural hazards allows for a better assessment of their risks and for a better preparation of the networks to reduce their impact. In Ref. [44], cloud computing is proposed as a platform to address the challenges faced by traditional ICT strategies for modeling the natural hazard processes. The higher capacity of computation and storage, along with its flexibility and scalability, allows for larger amount of data to be considered and, therefore, for more sophisticated models to be obtained. Cloud computing may be used in the process of disaster management (ways to handle the inaccessibility of cloud services during the disasters are also mentioned in Ref. [44]).

The analysis of the risks associated with different types of disasters has also been the subject of different studies. In Ref. [45], a survey of different methods of natural disaster risk analysis is presented. It is important to know: (i) the type of disasters that may occur; (ii) the likelihood of occurrence of disasters; (iii) the possible consequences of the disasters. Available data has to be systematically analyzed, so as to gain better knowledge on these three aspects and be able to prepare the infrastructures for disaster prevention and mitigation of its effects. In Ref. [46], common aspects of the risk assessment for different natural hazards are analyzed, taking into account different spatial scales (regional or continental).

The problem of the impact of natural hazards in critical infrastructures has been tackled by many researchers. In Ref. [47], the authors focus on the impact of hurricanes and floods on buried infrastructures, in particular water distribution systems. The impact of floods in critical infrastructures is also analyzed in Ref. [48], but from a perspective of recovery of a natural gas network. The problem of interdependence between different infrastructures (including communication networks) in pre-disaster hazard mitigation and post-disaster recovery is tackled in Ref. [49], with an application to the infrastructure with power, water, and telecommunication systems affected by a hurricane. Another important aspect in the post-disaster recovery efforts is the sequencing of repair actions on elements, as precedence constraints may greatly affect the final outcome [50]. A summary of the measures for risk minimization that may be implemented in the plants where the telecommunication or other supporting infrastructures are installed can be found in Ref. [20, Table 5].

The importance of the assessment of the risks associated to different natural hazards (tsunami, earthquakes, volcanic eruptions) that may affect a telecommunication network is stressed in Ref. [51]. The particular case of Hawaii, which is not only a bottleneck point in global telecommunication systems but is also affected by different large magnitude natural hazard events, is analyzed. The authors emphasize the importance of having solid information regarding past events, that may allow to estimate tsunami magnitudes and frequencies for Hawaii.

The challenges of submarine cable connectivity are mentioned in Ref. [52], as it is a structure prone to disasters. Still, it may be advantageous over satellite communications in terms of reliability, latency, and cost effectiveness, as well as being able to accommodate the increasing amount of traffic. This report exposes the vulnerabilities in all-optical submarine networks and mentions some ways to prepare for disasters, such as diverting the routes from areas prone to undersea earthquakes and having cables installed deeper in the sea near the coastal areas, to minimize the effect of storms.

A more recent paper dealing with the problem of disaster-aware submarine fiber-optic cable deployment is [53], which includes a proposal of providing a path pair (primary and backup cables). It follows from that paper that laying longer cables away from disaster-prone areas increases the deployment costs, but allows for much lower costs in the event of disasters. The impact of natural hazards in submarine fiber-optic cables is also explored in Ref. [54]. However, it is stressed in this paper that most failures are actually caused by human activities (e.g., fishing and shipping).

In Ref. [55], the use of wireless technology in environments prone to...
natural hazards (e.g., volcanoes or hurricane-affected regions) is explored. Due to the number and heterogeneity of wireless devices, an opportunistic routing paradigm is considered. Opportunistic routing is flexible and easily adapts to the network dynamics, which is particularly important in natural disaster recovery operations.

2.4. Malicious attacks in optical networks

In today’s data-driven world, where data networks are being used in every aspect of our life (from communications to the Internet of Things (IoT)), there is an ever-increasing need to protect the network not only from accidental failures but also from malicious attacks that can potentially target specific parts of the network infrastructure [56–60].

As optical networks provision connections with very high bit rates, even a short attack may have a very large impact on the information exchanged in the network. Thus, it is important that attacks are handled swiftly and decisively, so as to minimize their effect on the network.

Threats to the network include not only unauthorized access to data (eavesdropping attacks that target data confidentiality) but also attacks that have as an aim a disruption of the normal operation of the network and attack the availability of the network (e.g., denying service to the users of the network). As pointed out in Section 1 (and shown in Fig. 1), direct malicious attacks types that could potentially affect optical networks include EMP, DDoS, and physical destruction of nodes/links, whereas indirect attacks include cases such as destruction of the infrastructure due to bombings, etc. Clearly, attacks that physically destruct the infrastructure can also be viewed as massive failures/disasters and can be mitigated using techniques presented in this paper as well as in other works in the literature that deal with failure recovery (stemming from accidents, disasters, human error, malicious attacks, etc.). However, attack types that have effects similar to denial of service attacks, and affect the availability of the network, need to be viewed in a different manner. In general, to protect the availability of optical networks against service disruption and reduce the attack impact, either attack-aware provisioning techniques are utilized, or components are inserted in the network (equalizers, wavelength selective switches – WSSs) that can mitigate the spread of the attack. These techniques are discussed and analyzed in detail in Section 6.

The reader should note that there are several other types of attacks as well, including observing the existence of communications and thus violating the privacy of the communicating entities, attacks that focus on authentication services by the unauthorized use of network resources, as well as integrity attacks that are aimed at the destruction of data [61]. Optical layer security in general has received considerable attention from the research community and the network providers over the last few years, in an effort to develop different techniques and approaches (either during the network planning phase or the network operational phase) in order to prevent or mitigate such attacks [62]. The focus has been on jamming attacks (as described in Section 6) as well as on techniques to prevent eavesdropping attacks. For example, some works on eavesdropping via network coding techniques to provide confidentiality appear in Refs. [63–65], while other works on eavesdropping via spread spectrum techniques (such as Optical Code-Division Multiple-Access – OCDMA) appear in Refs. [66–73].

2.5. The need for post-disaster recovery

Post-disaster recovery of communication networks builds on the presumption that communication resources as well as other infrastructural resources will offer constrained resources, and an optimal post-disaster recovery scheme should ensure an optimal resource allocation strategy [74]. On the other hand, the challenge in the design of self-contained post-disaster recovery strategies is due to the uncertainty of the disaster scenarios. Therefore, post-disaster resilience strategies are beyond the survivable design of communication networks. For instance, in the case of failures in Passive Optical Networks (PONs) connecting mobile backhauls, a possible solution to retain the communication network active is to keep some of the radio access networks operational by using alternative power supplies including microgrids as proposed in Ref. [75]. Indeed, such a scheme would consider that the PONs connecting the backhauls cannot be powered up in a post-disaster recovery scenario but backup power would be available to support the radio access networks. It is worth to note that power outages are often a common consequence of various types of disasters [76].

The aftermath of catastrophic events is complicated and requires a thorough damage assessment/forecast through the aid of other networking technologies such as unmanned aerial vehicles (UAVs) [77]. Besides the need for self-contained aftermath forecast schemes, it is also crucial to equip the post-disaster recovery strategy with an emergency network solution such as a mobile cognitive radio base station as suggested in Ref. [78] to facilitate the communications between the networks connecting the affected population and the responders.

In addition to the aftermath forecast and the operation of the emergency networks, user-centric approaches are of paramount importance to model, predict and implement effective access solutions for high priority content in a post-disaster situation [79]. Furthermore, modeling the impact of a potential disaster can enable leveraging other network infrastructures in a post-disaster situation, which requires disaster-specific recovery solutions. An example to this concept is the design of a data communication network in the case of an earthquake [80]. Regardless of a disaster category, the communication in a post-disaster recovery is expected to be opportunistic; hence post-disaster recovery design of communication networks needs to incorporate the situation-aware service delivery [81]. Nevertheless, the advent of 5G technology and fiber-wireless convergence in the 5G ecosystem will facilitate the mitigation of post-disaster situation [82]. By leveraging the virtual reality, cloud computing and big data enabled in 5G networks, it should be possible to implement structural damage estimations, identify destruction levels and make effective online decisions for resource allocation.

Based on these observations, it should be underlined that the existing resilience approaches for communication networks require being coupled with effective, efficient and realistic post-disaster recovery solutions that care for emergency response, content evacuation, victim evacuation and service continuity.

3. Disaster resilience against earthquakes in optical wired networks

Earthquakes have devastating effects on telecom infrastructures. As few examples, in 2008 the Shichuan earthquake in China disrupted more than 30 000 km of optical-fiber cables and 4 000 telecom central offices [83], while the Great East Japan earthquake and tsunami in 2011 damaged almost 1 500 telecom offices on March 11 and 700 telecom buildings by the following aftershock on April 7 [84]. Given the importance and extensive impact of this type of disastrous failures on optical networks, in this section we concentrate on how to account for earthquakes in optical networks design. Specifically, we describe common modeling approaches, some selected techniques to prevent and counteract earthquakes’ impact on optical networks, and we outline possible future research directions.

3.1. Earthquake disaster modeling

An earthquake typically induces multiple spatially-correlated failures to the network infrastructure, mostly within the earthquake’s geographic footprint. The earthquake footprint varies according to the earthquake’s intensity, and to the characteristics of the affected territory (e.g., earthquakes in coastal areas induce tsunami-effects, while in mountain regions landslides can be originated). These failures can be also cascading (or, in other words, “time-correlated”), as earthquakes tend to hit in the form of seismic swarms, where several sequential shocks happen.
Given the multiple forms that an earthquake can take, it is important to develop generic modeling of the earthquake’s impact on an optical network. So, in the following, we start by discussing the modeling of the earthquake’s impact on optical (or more generically, telecom) networks, before providing an overview of the existing and under-investigation protection techniques to prevent the effect of these disasters.

There are two main existing categories of models for earthquakes in optical networks: deterministic models and probabilistic models.

**Deterministic Models.** A first simple, yet commonly employed, category of deterministic models is the one where network equipment (e.g., links, nodes, etc.) fail with probability 1 if they fall inside the geographic area where the intensity of the earthquake is above a certain threshold. This category of deterministic models is typically used as it is more easily tractable (both in simulation and in analysis) compared to the probabilistic ones. The most commonly-employed deterministic model for earthquakes is the SRLG (defined in Section 2.2) or Shared Risk Group (SRG). While SRLGs are lists of links that, for a specific threat defined beforehand (as, e.g., an earthquake), are expected to fail simultaneously, an SRG is a generalized list of network elements (containing not only links, but also nodes, interfaces, etc.). In order to predefine the location and the span of an earthquake to identify relevant SRLGs in a network graph, seismic hazard maps (see, e.g., the one in Ref. [85]) are used by network operators to gather information about the most risky earthquake zones and seismic hazard levels.

Fig. 4 shows a seismic map of the USA and a realistic long-distance optical network physical topology mapped over it. Existing studies from the field of seismology (see Ref. [86]) show that the minimum distance between a primary element (e.g., a link, or a node) and its backup resource to provide protection against earthquakes should be at least 50 km. These guidelines from seismologists help defining SRLGs, assuming that the most likely earthquake epicenters can be known. The resulting SRLGs are formed by all the elements belonging to a disaster zone, which are represented in Fig. 4 as purple circles.

The SRLG-based approach has been mostly applied on large (e.g., continental) networks, where only small sections of the network belong to highly seismic regions. In case of smaller networks (e.g., regional or national networks), we should change our initial hypothesis and assume that earthquakes may affect the entire network topology with approximately constant probability. In this case, tools for disaster modeling come from the field of computational geometry [87]. Computational geometry can be used to identify the most vulnerable part of the network, i.e., the locations where an earthquake could have the most catastrophic effect on the network. To leverage tools and algorithms from computational geometry, the region impacted by the earthquake is typically represented either by a line segment cut or by a circular cut that removes all links/nodes that intersect it [88]. As an example, Fig. 5 shows two potential disaster regions, L and C. Disaster zone L is represented as a line segment, and two links are damaged when an earthquake occurs. Another potential disaster covers the circular area represented by region C and damages one node and four links when it occurs.

These computational-geometry-based approaches can be used to gain insights on network vulnerability to multiple correlated failures, as those in seismic swarms. In Ref. [88], authors proposed a method to determine the number of earthquakes, modeled as circular disks, that disconnect a pair of nodes. This information can be used to design the earthquake-survivable optical network.

**Probabilistic Models.** While a deterministic model assumes that all network elements within a disaster zone entirely cease to work, in practice, network equipment fails with a certain probability which depends on many factors, such as the equipment’s size and manufacturing characteristics (i.e., impact and vibration resistance), and its distance from the earthquake epicenter. So, for example, as the intensity of an earthquake attenuates for increasing distance from the disaster epicenter, a probabilistic failure model [89] should consider that a fiber duct fails with a probability which depends on the duct length and on its intersection with the earthquake footprint. Hence, to calculate these probabilities, the optical network can be divided into a grid of possible “failure regions” which are then divided into a set of consecutive annuli (of equal width), and then assume that a fiber duct falling within an annulus fails with a decreasing probability. Similarly, in Ref. [32], authors suggest to define a super-level set, i.e., the set of locations for which the probability of equipment failure is greater than a certain value. As for the deterministic models, also probabilistic models can leverage seismic hazard maps as the one shown in Fig. 4. Assuming a predefined set of disasters (i.e., SRLGs), a formula to estimate the probabilistic risk of an earthquake affecting a network can be derived as proposed in Ref. [90]. Ref. [91] derives a model for earthquake risk on backbone optical network (the “ERBON” model) that joins contributions from stochastic distributions, real statistics, stochastic geometry, and graph theory to model seismic zones.

### 3.2. Description of selected schemes

In this section, we survey techniques to provide resilience against earthquakes in optical networks that are broadly classified into proactive and reactive approaches [92]. In proactive approaches, disasters are assumed to happen in deterministic (or probabilistically) modeled in advance, and connection requests have to be provisioned a priori such that the disaster effects are minimized. In a reactive approach, the actual disruption due to the earthquake is given, and connections are re-provisioned after the earthquake to restore the maximum amount of traffic possible. In some cases, connections already established may release some network resources by accepting a certain degree of degradation (i.e., degraded-service tolerence). The objective is to serve as many connections as possible with an acceptable quality of service in an extreme situation, as, in the aftermath of a disaster, network congestion (also known as “traffic crunch”) can be experienced.

**Disaster-Aware Provisioning and Reprovisioning.** One traditional proactive approach is to provide a pair of primary and backup lightpaths for each end-to-end connection, in such a way that the primary and backup lightpaths are routed in two distinct SRGs (i.e., SRG-disjoint). In this way, the two paths are not simultaneously disrupted in the case of an earthquake. However, as this approach may require very high bandwidth over-provisioning (and hence become economically unsustainable), the concept of Disaster-Aware Provisioning (DAP) [93,94] has been introduced. To enable DAP, first the no/low-risk regions of the network are identified. After that, the most important connections are routed along these regions to lower the risk of disruptions (and the consequent penalties due to data loss) in case a disaster occurs, and the routing is not necessarily protected with a backup path. In case the primary connection is disrupted, the reconfiguration capability of modern network equipment can be exploited to re-provision some or all connections. This technique, called Reprovisioning (RP), is applied not only to connections in regions directly affected by the first earthquake shock, but also to connections under risk of correlated cascading failures due to secondary shocks [95].

Fig. 6 shows an example of disaster-aware provisioning where n1 and n2 are two pre-determined earthquake zones. Let us consider that if earthquakes in n1 and n2 occur, network equipment within n1 and n2 fails with probability of 0.2p and 0.5p, respectively. We consider three connection requests: (1) t1 from node 1 to 9; (2) t2 from node 2 to 8; and (3) t3 from node 1 to 3. In Fig. 6(a), the connection requests from t1 to t3 are routed on shortest-path lightpaths; in this case, if earthquakes in n1 and n2 occur, all three connections are disrupted. In Fig. 6(b), connection requests are routed using the disaster-aware provisioning approach. Namely, if earthquakes in n1 and n2 happen, only t2 is disrupted since t2’s destination is within n2. Overall, the penalty caused by n1 is totally alleviated, while the penalty caused by n2 is significantly reduced. The main intuition of DAP is that, in case of rare events such as earthquakes, it is too costly to deploy a full protection against these events, hence probabilistic approaches that minimize the risk of being affected by a
disaster aim to at least minimize the impact/disruption of earthquakes.

**Manycasting, Multipath Provisioning and Degraded Service.** Future networks (e.g., the incoming 5th generation of mobile communications) will support a variety of services with very diverse requirements in terms of bandwidth, latency, and reliability. To cope with such diverse requirements, using the same protection scheme for all services would result in suboptimal solutions (either too costly, or not suitable for ultra reliability services as, e.g., those in the area of autonomous driving and industrial automation). Thus, the different tolerances to the disruption of different services must be leveraged to tailor survivability techniques to the specific service requirements. For instance, some applications like video streaming or file transferring can still be operated with a reduced bandwidth or a lower resolution if a widespread network failure prevents from delivering the full service. The concept of **degraded service (DS)** refers to the possibility of delivering the reduced amount of network resources for a service vs. a normal operation requirement, and is attracting considerable attention. DS provides the opportunity to reduce network disruption, protection cost, and increase the maximal number of admitted connections in the case of a disaster.

To enable degraded services, the concept of **partial protection** (also known as bandwidth squeezed protection in the context of elastic optical networks – EONs) can be leveraged [96]. In a traditional dedicated protection, network resources are reserved in advance for both a primary and a backup lightpath such that, if the primary lightpath is disrupted, 100% of its bandwidth is guaranteed (i.e., full protection). Considering earthquakes are statistically rare events, 100% over-provisioning may not be practical. Providing protection for a portion of the requested bandwidth, or partial protection, can eliminate the extensive resource reservation characteristic to full protection schemes [97]. The fraction of the requested bandwidth that will be ensured, even under severe failures, is computed according to the degradation tolerance of the services, generally stipulated in a Service Level Agreement (SLA) between customers and network operators.

Moreover, the shifting paradigm toward cloud computing provides new opportunities for survivable provisioning against earthquakes. In cloud networks, services are replicated at multiple DCs from which users gain access to the service content. This virtualization of the actual location of a service gives users an opportunity to access the service at any of the locations where the service is replicated. In terms of routing/resource assignment in optical networks, this means that the advanced routing paradigms such as anycasting and manycasting can be used. In anycasting, a user can be served by any of the DCs that host the requested service. In manycasting, a user is served by any subset of the DCs that host
the requested service. These advanced routing paradigms create unique opportunities to improve the resilience of cloud services.

Fig. 7 shows three different provisioning schemes to illustrate the concepts of partial protection and anycasting on a 7-node physical network with DCs deployed at nodes 5, 6 and 7 [96,98]. The optical network contains three Disaster Zones (DZs) D1, D2, and D3. We consider a service with 10 Gbps required bandwidth from node 1 and assume all DCs have the requested content. Fig. 7(a) shows a proactive, full protection provisioning scheme where a backup lightpath is established to DC at node 7, which is SRG-disjoint to the primary lightpath to DC at node 6. When either D1 or D2 occurs, this scheme still fully provides 10 Gbps bandwidth but it has consumed 20 Gbps bandwidth reservation. Even with such high network resource consumption, in the worst case when two disasters D1 and D2 happen simultaneously, node 1 cannot access the service anymore. Fig. 7(b) represents multipath provisioning destined to a single DC (i.e., node 6). In this case, even if disasters D2 and D3 occur, service for node 1 is still fully protected even if the overall bandwidth reservation is 15 Gbps (in general multipath provisioning allows to save backup capacity). However, this provisioning scheme is vulnerable to failures caused by disaster D1 since node 6 (the destination DC) is within D1.

To overcome this weakness, multipath provisioning to multiple DCs (i.e., manycasting) can be employed, as shown in Fig. 7(c). The traffic request from node 1 is split and routed to DCs at nodes 5, 6, and 7 with equal bandwidth reservations (i.e., 5 Gbps). In total, this provisioning scheme still consumes 15 Gbps bandwidth, while it provides significant improvement of survivability. In this scenario, full protection is guaranteed after any single disaster. Also, partial protection is ensured with at least 5 Gbps bandwidth to node 1 (degraded service) after two random disasters. Note that this proposed scheme is also robust against node/DC failures. In a large network, a request from a cloud user is served by a subset of DCs hosting the desired content. We refer to this provisioning scheme as manycasting. This illustrative example shows the potential of multipath manycasting provisioning for cloud services to improve survivability.

3.3. A critical comparison

In Table 1, we provide a summarized critical comparison of the resilience techniques against earthquakes in optical networks mentioned above.

We classify DAP, Multipath and (M)Anycasting as proactive provisioning and RP as reactive provisioning. In general, proactive techniques are inflexible but have shorter recovery time, while reactive techniques are flexible but have longer recovery time. DAP is the solution that offers lower cost owing to its capability to leverage earthquake-risk-minimized routing, but it may lead to long paths, as well as unbalanced network resource utilization. RP is the only solution that offers adaptability to any type of failure extent, but at the price of a “restoration-type” recovery time. Multipath is a promising technique in terms of resource savings, whose adoption is limited mostly by the necessity to compensate for the differential delay at the destination. Finally, manycast- and anycast-based routing have the great property of defending against the occurrence of DC failures (quite likely in an earthquake scenario), with replication of service locations (destinations) that can be exploited to provision static services such as file transfers or media streaming, while it entails complex synchronization issues in the case of more dynamic services (e.g., social network platforms).

Also, even if not explicitly mentioned in Table 1, accounting for the flexibility provided by degraded service provides a huge potential to save network resources and improve its survivability.

1 Some research, for example [99], has showed that multipath provisioning can effectively be applied even in multi-rate optical networks.

3.4. Future directions for earthquake resilience

With the advent of 5G communications, the focus of network operators is moving towards the metro-access segment of the telecommunication infrastructure. In fact, metro networks are currently evolving from a rigid aggregation infrastructure to a composite network-and-computing ecosystem supporting new critical services (as, e.g., autonomous driving). This evolution is enabled by several new technical advances such as Software Defined Networking (SDN), optical/wireless convergence, edge computing, and network slicing. All these new technical advances represent the building blocks to create new resilience strategies against earthquake disruptions.

As already mentioned, today’s metro networks are edge cloud networks where popular contents/services are hosted in DCs close to the end users (see, e.g., the concept of Central Office Reconfigured as a Datacenter, or CORD [100], being developed for traffic offloading and/or latency reduction). In these networks, traditional approaches to network survivability should be updated to reflect the evolving reliability requirement. Most 5G critical services can be provided as long as a replica of the service content is available in each disconnected network partition, even under a large-scale disaster such as an earthquake.

To adapt to this trend, future protection approaches should consider evolving the traditional survivability metric, namely network connectivity, and investigate how it can be evolved towards a new concept, called content connectivity. Let us refer, as an example, to the concept of Survivable Virtual Network Mapping (SVNM) with content connectivity to introduce this new concept [101].

Network Connectivity (NC) is defined as the reachability of any network node from all other nodes in a network. Originally, NC has been used to measure network survivability in end-to-end communications and will probably remain the default option for a smaller failure scenario (such as a random single-link/node failure). Unfortunately, in the case of a large-scale disaster, multiple links and nodes may be simultaneously interrupted, and ensuring NC can be very costly, or even infeasible.

Content Connectivity (CC) is defined as the reachability of the content from every node in a network under a certain failure scenario. The main idea is that, even if the network becomes disconnected, as long as CC is guaranteed, every user can still reach at least a content replica in all disconnected network partitions. Therefore, service continuity is guaranteed.

To visualize the impact on survivability of the content connectivity concept, let us consider a survivable virtual network mapping (SVNM) example of a 4-node virtual network (VN) from Fig. 8(a) over a 6-node physical network (Fig. 8(b)). In the physical network, nodes 1 and 4 host two content replicas of interest for the VN. In Fig. 8(c), the VN is mapped over the physical network such that no physical link supports more than one virtual link. Thus, the VN is network-connected survivable (i.e., NC survivable) against a random single-link failure in the physical network. That is, there is no single link in the physical network whose removal disconnects the VN. In Fig. 8(d), the VN is mapped over the physical network in a way that the physical link (1, 6) carries both virtual links (1, 5) and (1, 6), and its failure would disrupt NC for the VN. Nonetheless, content replicas at nodes 1 and 4 remain accessible to every virtual node. Hence, the VN is CC-survivable. Note that:

1. the mapping in Fig. 8(d) utilizes one physical link less than the mapping in Fig. 8(c); and
2. the mapping in Fig. 8(d) is survivable in the higher layer (e.g., IP layer, i.e., node 1 can use node 5 as a transit node to reach the content replica at node 4) rather than in the optical layer.

Numerical results show that mapping with content connectivity can guarantee a higher service survivability, especially in the extreme multiple failure situations such as earthquakes, and provide a significant improvement of service availability.

While most of the discussion so far has concentrated on proactive
solutions for the pre-disaster phase (preparedness), interesting directions are arising also regarding the post-disaster phase. A classical problem in the post-disaster phase is the progressive network recovery (PNR) problem [79,102], which consists of finding the best sequence of links to be repaired in the case of multiple links failure in order to maximize the cumulative traffic carried during the recovery process. Today, especially in metro areas, operators can employ portable professional equipment (e.g., portable amplifiers, recovery truck containing the movable base stations or movable DCs), and the traditional PNR problem must be evolved to consider how to schedule a rapid recovery via deployable recovery units transported over specialized trucks that can restore the network connectivity or computing facilities [103].

4. Preparedness for foreseeable disasters

Some natural disasters may be foreseeable, even if the full extent of their impact (in terms of affected area and the intensity of the damages) may not be completely known in advance. Foreseeable disasters include tornadoes or heavy rain, that may cause floods and landslides, and also volcano eruptions, as they are tracked by weather and geological organizations. For predictable disasters, some procedures may be executed in advance, so as to minimize the impact of the disasters in the communities (people and infrastructures). In fact, some foreseeable events have severely damaged the communication infrastructure and affected disaster recovery. Due to hurricane Katrina, communications failed totally for the City of New Orleans in 2005, as a result of the strong winds and flooding that followed, strongly hindering the response of emergency service organizations [104]. Namely, over 3 million customer phone lines were out of service in the areas of Louisiana, Mississippi and Alabama [105]. Close to 100 broadcast stations failed to transmit, and hundreds of thousands of cable customers lost service. Hurricane Sandy in 2012 caused 300 Verizon wired centrals to be affected by power outages, but only two key central offices in Lower Manhattan, where the flood affected their backup generators [106], did fail. The impact was much more significant on wireless communications, where an average of about 25% of the base stations in the affected area lost service [106]. More recently, hurricane Michael affected Florida in 2018, interrupting 40 000 communication lines [107].

In the remaining part of this section, we point out the need for disaster modeling and efforts in that field as they allow to produce the
forecast alerts for predictable disasters. Next, we describe and compare data evacuation and VM migration strategies which allow to reduce the impact of predictable disasters in data-centric networks. Finally, some future research directions are outlined.

4.1. Disaster modeling

The problem of floods modeling and impact assessment is tackled in Ref. [108]. Machine learning methods may be used not only to process remote sensing data to estimate the impact of floods, but also to process the social media data to tackle the response to the floods. Therefore, it proves to be useful in preparation stages (development of mitigation and response plans) and in the aftermath of the events (development of recovery plans). Knowledge of the areas which are most vulnerable to floods is quite important to devise the appropriate preparedness and response plans. An example is Fig. 9 taken from Ref. [109], where the social media data is used to predict the flood impact.

In Ref. [111], the assessment of flood risks is analyzed in terms of the spatio-temporal dynamics of the processes that may originate floods, which include heavy rain and river overflows. The assessment of the risks due to heavy rainfall is tackled in Ref. [112]. A statistical analysis is performed on available data from the past years, allowing to obtain a conceptual model of the hazard. Flood modeling is accomplished in Ref. [113] by means of two complementary methods: existing flood inundation models and satellite flood data. This way, in the particular situations of coastal areas where the flooding model uncertainty is high, the satellite data may help in defining the values of some model parameters and thus improve the quality of the model.

![Fig. 9. Map of Tuscany, Italy including information on orography and areas prone to floods (in blue) and to landslides (in red); the location of 27 schools in the region is also displayed [109, Fig. 1).](image)

In Ref. [114], a review of disruptions and physical damages to elements in critical infrastructures after volcanic eruptions, is performed. Four different volcanic hazards (tephra fall, Pyroclastic Density Currents (PDC), lava flows, and lahars) are considered and their effects on different critical infrastructures are analyzed. In particular, in terms of the impact on communication networks the authors identify two major problems: (i) the overload of the network due to an increased demand for services accompanied by some signal attenuation and interference (due to the tephra falls, PDCs and lahars); (ii) the destruction of network elements, including lines, cabinets, or exchanges (due to lava flows). The authors establish different levels of impact for each of the considered volcanic hazards.

The assessment of damages caused by hurricanes in communication networks is described in Refs. [115,116]. Hurricanes and tsunamis may originate landslides and floods, which may damage power backup equipment. In fact, according to this study, most issues in communication networks derive from the lack of power, i.e., power outages affect communication equipment and lead to service disruptions. Preventive measures to account for the possibility of floods include the appropriate location of power equipment and/or the use of waterproof doors in the facilities. Renewable energy sources and making sure that there is enough fuel for backup generators, may also help in dealing with the foreseen power outages in these situations [20]. The papers [115,116] also discuss some network vulnerabilities due to the use of fiber-optic Remote Terminals (RTs), as they need to be powered. Typically, backup batteries last for about 8 h, which may be insufficient in many cases and ends up providing a worse service than copper multi-pair cables. The downside of using copper cables is the likelihood of corrosion due to water; the use of water pump systems is usually not possible, again due to the lack of power.

Fires can also be considered as foreseeable disasters and they are known to cause serious damage to optical networks, both in aerial cables and in buried cables at shallow depth, as well as in the towers and poles supporting telecommunication lines [117]. The estimation of the so-called dead fuel conditions, along with the information on meteorological and environmental data, may be used to predict the occurrence of fires and how they will spread [118]. This information may also be used to predict the most risky forest fire areas, allowing for the update of maps of probability of extreme events. Nevertheless, the accurate prediction of time and place for a fire remains elusive. Moreover, the progression of fires is more difficult to predict than other natural disasters as, e.g., a change in the pattern of winds is sufficient for the fire to spread in an unexpected way. There are numerous works on models on wildfire spread prediction and its dependence on the accuracy of the required input parameters, such as [119]. Some preventive measures are put forward in Ref. [117], such as the materials to be used in the poles and the clearance around them.

4.2. Description of the selected approaches to mitigate the effects of foreseeable disasters

For some weather-based disaster situations (e.g., hurricanes, floods, or tornadoes), it may be possible to issue forecast alerts, that will allow for some preventive measures to take place in order to mitigate the effects of the disasters. Artificial Intelligence (AI) and IoT have been extensively used in this context [120]. The problem of the migration of data (in particular, critical data) between DCs when a disaster alert is issued, has been tackled in different papers. As mentioned in Section 3.2, the possibility of replicating information at multiple DCs and the possibility of virtualizing services in different locations offer numerous opportunities for survivable provisioning against disasters, in particular foreseeable disasters. Note that data evacuation may also be accomplished in some post-disaster scenarios, as discussed in Section 7.3.2.

In Ref. [121], the authors propose to avoid the effects of heavy rain by reconfiguring a logical network according to weather information. This reconfiguration includes the migration of VMs, to avoid disasters.
The paper tackles issues such as the applicability of this idea to large networks, the parameter adjustments or the effectiveness of this process when compared to the simple reconfiguration of routes. In Ref. [122], the problem of the migration of VMs pertaining to DCs that may be impacted by predictable disasters, is tackled. In order to guarantee that during the migration (almost) no service interruptions will occur, an ILP model aiming at the joint maximization of the number of migrated VMs and the minimization of the service downtime, the occupation of network resources and duration of the migration, is put forward. These different criteria may be conflicting objectives, so a trade-off between them is explored. Experimentation with online and offline strategies for the VMs migration in a reference optical network is presented. The online migration process has several iterations as can be seen in Fig. 10. The first iteration copies the main memory of the VM and subsequent iterations (separated by a network delay $\tau$) depend on the amount of modified memory during the previous iteration. The procedure ends after a certain number of iterations or if the amount of the modified memory is below some threshold. This process is possible due to the alert time before the disaster impacts the VM being migrated.

In Ref. [123], a heuristics selects the paths for data evacuation in an anycast network model, such that the delay for that evacuation is minimal. To calculate the delay, different parameters are considered, such as the propagation delay, network bandwidth, and congestion. An evacuation deadline may be established, but depending on the type of a disaster, it can range from milliseconds to minutes or even longer. The goal is to evacuate the maximum content in the established evacuation deadline, or a certain amount of content in minimum time. A follow-up to this work is [124], where the possibility of other node/link failures along the routes used for data evacuation is considered, which may increase the extent of data losses.

Another approach [125] starts by defining parameters that reflect the network service performance, in terms of bandwidth, response time, delay, or safety requirements. Given this information, zone risks in DC networks may be defined. Data should be placed preferentially in DC nodes in the zones of the lowest risk, so that when there are alerts for a disaster in the area, the amount of data that needs to be evacuated is minimal. With this approach, only a small amount of data has to be evacuated before the disaster, which increases the efficiency of the evacuation process. Note that this strategy is implemented in a DC network based on SDN, taking advantage of the centralized control logic, that allows for an allocation of resources according to the performance requirements. In Ref. [126], the amount of data to be transferred in the event of a disaster is also reduced, as copies of blocks of data are regularly kept in backup sites during the usual VM operation.

In the context of disaster preparedness, the problem of the location/placement of backup DCs is tackled in Ref. [127]. Backup DCs must be placed in low risk locations, but still close to the main DC, so that the data may be migrated to the backup facilities quickly. A multi-objective problem is formulated, taking into account the expected disaster loss and the evacuation latency. The resolution of the problem is accomplished by the Ant Colony Optimization (ACO) metaheuristic approach.

Earlier papers on this subject are [128,129], where the problem of efficient distribution of storage resources (disks) and computing resources (VMs) in network nodes is tackled. Both remote storage needs and virtualization techniques have to be taken into account. The allocation of resources is such that two conflicting aspects regarding the backup facilities are considered: (i) they should be close to the main DC to minimize the delay in accessing the information and the network load; (ii) if a node fails, the VMs should migrate in a swift way and resume functioning in those facilities using the stored backup data. The purpose is to minimize the impact in the performance of the different applications. In Ref. [128], every VM must be associated with two disks (a local one and a remote one, with a backup copy of the application data), and one disk should be assigned to one VM only. Different ILP formulations for optimization problems are proposed:

- (i) the problem of the minimization of the maximum bandwidth utilization on the used network links, which allows to control the network congestion and to limit the maximum delay between VMs and the remote backup facilities;
- (ii) the problem of the minimization of the total number of hops between VMs and their backup disks, again to limit the maximum delay between VMs and the remote backup facilities;
- (iii) in the event of VM migration due to a disaster, it is important to guarantee that the site to where the VM migrates (usually the site where its backup disk is located) still manages to keep the normal operation for the VMs that were already there – this is accomplished by minimizing the number of new VMs in a given site;
- (iv) the same problem as the previous one, but with an additional constraint on the maximum number of hops between the VM location and the related backup site.

Heuristic approaches are used to solve these problems. In Ref. [129], different ILP problems are put forward:

- (i) the Network-Aware Problem (NAP), aiming at the minimization of the maximum (or the mean) number of hops between each VM and its backup disk, so that the end-to-end delay between VMs and backup facilities is low;
- (ii) the Disaster Recovery Problem (DRP), aiming at the minimization of the overload on backup sites that will host the migrated VMs in the event of disasters;
- (iii) hybrid problems, Disaster Recovery Constrained (DRC)-NAP and Network Constrained (NC)-DRP.

These latter problems are bi-objective and are solved in a lexicographical way: in the DRC-NAP problem, the DRP is solved and among all the optimal solutions, the one leading to the minimization of the maximum (or the mean) number of hops between each VM and its backup disk is selected; in the NC-DRP, the NAP is solved and among all the optimal solutions, the one leading to the minimization of the overload on backup sites that will host the migrated VMs in the event of disasters is selected.

4.3. A critical comparison

In Ref. [130], the resilience of ICT infrastructures in the event of disasters is analyzed. Based on the assessment made by different authorities in the aftermath of disasters, two main conclusions arise: (i) considering guided transmission, buried cables are preferable to aerial cables: although the former may be affected by floods and earthquakes, ducts with buried cables tend to be more resilient than the aerial cables (whose poles are subject to collapses in situations of extreme winds); (ii) hybrid ring-mesh architectures prove to be more reliable than ring topologies (very susceptible in the case of multiple network failures) and more cost-efficient than mesh network topologies. These conclusions are reinforced in other studies (e.g., in Ref. [29]), along with other

---

**Fig. 10.** Illustrating the iterative nature of the copy phase during a VM migration (adapted from [122, Fig. 1]).
recommendations, namely the implementation of adequate power sources (backup and/or use of renewable sources), the geographical distribution of the Communication Systems (CSs), the diversification of the used CSs and the establishment of rules that may help in managing the network congestion.

Regarding the pre-planned data evacuation and VM migration strategies, three major possibilities are proposed in the literature (see Table 2):

- Given the existing placement of the backup facilities and the evacuation deadline (i.e., the estimated time until the disaster actually strikes the area, which may range from the order of milliseconds up to a few hours, depending on the nature and the severity of the disaster), the aim is to evacuate all the data or at least the maximum amount of data until that established deadline. The migration of VMs is done in the same way, i.e., if not all may be migrated, the attempt will be to migrate as many as possible, so as to minimize the downtime.
- In some references, different zones are analyzed and classified in terms of their risk of disaster impact. According to that classification, the most crucial data are kept in nodes with a low risk. This strategy has the disadvantage that data may be kept in nodes far from the places where they are most needed, which presents problems in terms of delay and ease of access. The advantage is the safeguard of the most critical data, and also the reduction of the amount of data to be evacuated. It can be used in conjunction with the previously described strategy.
- The amount of data to be evacuated may also be reduced by a proactive cache of information at backup sites. Different references focus on the appropriate placement of backup DCs, which ideally should be located near the main DCs but distant enough from the areas with higher potential of suffering from damages in the event of disasters.

Regarding the VMs migration, it is preferably done to the DC containing the backup information. Note that it is important to guarantee that the site to where a VM migrates should be able to accommodate that new VM, and still continue to provide the services to the already existing VMs in that location.

### 4.4. Future research directions

The integration of SDN in optical networks – Software Defined Optical Networks (SDONs) – may have an impact on the preparedness of optical networks for disasters. The global view of the network topology, the switches and the information on performance metrics (bandwidth or latency of links) provided by the centralized controller in a software defined network, is very important for the network-aware VM placement and the traffic management, in particular related to VM migration. For this purpose, controllers must be adequately located in the network. In Ref. [131], a dynamic method for optical controller placement is put forward, taking into account different optical controllers, resource limitations, latency requirements and costs.

The problem of VM migration in SDN-enabled cloud DCs is tackled in Refs. [132,133]. In Ref. [132], the performance of the migration in terms of migration time, downtime, and the amount of transferred data is considered. In Ref. [133], the proposed approach is based on the exploration of the network information over time to dynamically calculate the network-wide communication cost of the traffic flows. This way, it is possible to devise when the migration is appropriate. In the event of disasters, this approach would have to be adapted, as the question is not when to perform the migration, but rather how to perform it in the most efficient way. Still, the advantages of SDN regarding the monitoring and the reaction to real-time changes may prove to be very important when disasters occur.

Examples of other approaches for VM migration in the context of SDN are described in Refs. [134,135] and in Ref. [125], as mentioned previously.

### 5. Performance of optical wireless networks in the event of weather-induced massive disruptions

The effects of weather disruptions on the performance of optical wireless networks are addressed in this section. Example issues include, e.g., reduced capacity of optical wireless links due to the changing weather conditions such as fog or snow. Although such effects are temporal and do not lead to permanent failures, due to their frequency and the possibility to occur at multiple locations at a time, mitigation of their impact to sustain the network performance is similarly important.

OWC systems using Light Amplification by Stimulated Emission of Radiation (LASER) or Light Emitting Diode (LED) generating the optical signal (visible light communications – VLC, infrared – IR, and ultraviolet – UV [136]) in the wireless domain at wavelengths in the 780–1600 nm range can be viewed as an important alternative or a complementary solution to fiber wired infrastructures especially in the access areas of 5G and beyond installations [137,138]. They are also seen as an important “last-mile” connectivity concept for metropolitan networks [139], as well as point-to-point and point-to-multipoint configurations [6] increasing the capacity of other existing networks.

With the unlicensed band, the link capacity of several Gbps and bidirectional (full-duplex) transmission using different wavelengths in each direction [7], OWC links are also likely to replace in many areas the radio frequency (RF) links offering the capacity of at most several hundred Mbps [138]. Following [140,141], an OWC link established using a laser at 200 THz can provide capacity being almost 200 000 times greater than a 2 GHz RF link. Experimental configurations have also shown the potential of OWC systems to achieve even higher data rates of up to 1.72 Tbps over a distance of more than 10 km [7,142] during good weather periods.

Although OWC technology is meant to be utilized for communications between stationary nodes, it can also take part in scenarios involving moving objects [7]. Following [6], a full list of possible use cases comprises:

- ground-to-ground configuration including short- and long-distance terrestrial links;
- satellite uplink/downlink;
- inter-satellite connectivity;
- deep-space probes to ground transmission;
- ground-to-air/air-to-ground communications (for instance UAVs operating at short distances at a lower altitude below the cloud layer,

### Table 2

Pre-planned data evacuation strategies.

<table>
<thead>
<tr>
<th>Strategies</th>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evacuate max data in min time [123,124]</td>
<td>Lower delay for access</td>
<td>Dependency on the residual capacity of the links</td>
</tr>
<tr>
<td></td>
<td>Ease of access</td>
<td></td>
</tr>
<tr>
<td>Placement of critical data in low risk zones [125]</td>
<td>Lower amount of data to be evacuated</td>
<td>Increased delay to access the data</td>
</tr>
<tr>
<td></td>
<td>Efficiency of the evacuation process</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Safeguard of critical data</td>
<td></td>
</tr>
<tr>
<td>Regular copies of data in appropriately placed backup DCs [126-129]</td>
<td>Lower amount of data to be evacuated</td>
<td>Overload on backup sites</td>
</tr>
<tr>
<td></td>
<td>Efficiency of the evacuation process</td>
<td></td>
</tr>
</tbody>
</table>
or long-distance high altitude platforms (HAP) above the cloud layer).

OWC solutions can be utilized in networks with distances ranging from several meters up to thousands of kilometers. In particular, apart from the ultra-short distance communications (for instance chip-to-chip communications) or short-range solutions (e.g., concerning wireless body area networks – WBANs), as presented in Refs. [136,140], OWC solutions can be further classified into three main types:

- optical wireless home networks (OWHNs) often referred to as indoor IR/VLC optical wireless home networks offering broadband communications inside buildings. OWHNs support transmission rates typically of several hundred Mbps and are meant to serve as local area networks consisting of cells in each defined space in the building. In each cell, there is a base station to which the end-user terminals are connected via short-range infrared and LED links;
- long-range optical wireless terrestrial networks (OWTNs) also called Free Space Optical (FSO) networks [18] with point-to-point outdoor links via the line-of-sight (LOS) outdoor channels characterized by atmospheric turbulence. Depending on the environmental conditions the distance between two transceivers forming an FSO link in OWTNs ranges from several hundred meters to several kilometers. The major use scenarios of OWTNs include: bridging the existing networks with distances ranging from several meters up to thousands of kilometers. The major use scenarios of OWTNs include: bridging the existing geographically separated wired networks; last mile high-bandwidth connectivity; integration with wireless RF networks to overcome their limits (mainly throughput);
- ultra-long range optical wireless satellite networks (OWSNs).

A configuration example of the outdoor OWC equipment and its integration with the existing wired infrastructure is presented in Fig. 11.

In the remaining part of this section, we primarily focus on the long-range outdoor terrestrial FSO links due to the advantages of their use in contemporary communication network architectures that include electromagnetic compatibility (EMC)/electromagnetic immunity (EMI), and security aspects [5,140]. It is also important to mention their low cost of deployment as opposed to wired networks. In particular, for FSO installations, there is no need to install any cables (e.g., underground). Optical wireless components are also less expensive than their respective RF counterparts and consume less power [140]. Additionally, as the optical spectrum is license-free, obtaining the permissions to use certain optical channels is not needed. Another advantage is that the point-to-point and narrow beam FSO transmission is characterized by a low probability of interception (LPI)/low probability of detection (LPD) and negligible mutual interference [140]. A variety of FSO equipment offering the transmission rates of 10 Gbps or higher is already available on the market [136].

All advantages of FSO equipment described above justify their usability by a wide range of applications including, e.g., broadcasting of live events, back-haul for cellular systems, serving as redundant links/topologies, enterprise connectivity, or video monitoring [136].

For a complete picture of FSO characteristics, when designing a robust communication network involving at least a partial utilization of FSO technology, it is, however, important not to forget about its drawbacks related to communication resilience under certain conditions, as presented in the following part of this section.

### 5.1. Challenges to resilient FSO communications

FSO installations face a number of random challenges [140,143], shown in Fig. 12, including:

- the adequate pointing, positioning and tracking of signal due to tracking system errors or problems with mechanical alignment (which is not simple in optical wireless networks even for stationary nodes);
- Mie scattering resulting in the deflection of part of the light beam away from the assumed receiver;
- the atmospheric turbulence affecting the signal propagation (due to random changes of the atmospheric refractive index) leading to the degradation of signal-to-noise ratio (SNR) and increased bit error rate (BER) even up to the level considered as a link failure;
- weather conditions, especially the impact of fog, clouds, and snow.

In particular, concerning the impact of adverse weather factors, fog has been confirmed to cause the greatest losses in FSO link capacity among all weather conditions (with the attenuation often considerably over 30 dB/km [6]). Due to the size of water droplets comparable to the size of the infrared FSO wavelengths [143], a substantial Mie scattering of the laser energy is encountered especially in periods of thick fog [6]. Following [7], Mie scattering apart from its presence due to dense fog, can also appear due to clouds – especially in the case of the space-to-ground FSO links.

It is important to note that the impact of fog is very much dependent on its characteristics. In particular, two types of fog are possible: maritime and continental fog. Maritime fog is characterized by a considerably high attenuation of up to several hundred dB/km (similar to clouds), while the attenuation of the continental fog was found to be lower – about 100 dB/km [7]. Also, continental fog was found to be more stable than maritime fog that exhibits frequent fluctuations of attenuation.

For FSO networks, the impact of snow has been shown to be most often moderate (positioned between the impact of light rain and a moderate fog [136]). Unlike the RFs – especially the millimetre-wave frequencies in the 30–300 GHz range – which are much impacted by

---

**Fig. 11.** Example deployment of the OWC equipment and its integration with the optical wired infrastructure.

**Fig. 12.** Challenges to resilient FSO communications.
Hybrid RF/FSO Architectures. Since the weather-related events affecting FSO and RF systems are diverse (for instance, FSO links are sensitive primarily to fog, snow, and clouds, while RF links suffer from the impact of rain), these two technologies are often considered complementary (see Table 3). This, in turn, explains the existence of hybrid RF/FSO systems with parallel FSO and RF links installed between each pair of neighboring nodes. RF/FSO systems, by default, use both types of links under normal weather conditions. Their high level of resilience to weather disruptions is assured, as in the case of adverse weather conditions at least one type of links remains operational (i.e., RF links in periods of fog/snow/clouds and FSO links in rainy periods) [7,136].

It is important to note that the complementarity of RF and FSO technologies also refers to capacity and coverage. In particular, RF systems have been shown to be able to diffract, reflect, penetrate, and scatter on the obstacles, as well to easily maintain connectivity with neighboring nodes. All these characteristics of RF systems come at the price of reduced capacity, operation at licensed frequencies, and increased risk of channel interference. FSO systems, in turn, offer communications at license-free frequencies and at much higher data rates, as well as at practically non-existing risk of signal interference; nevertheless, they require LOS communications and a precise setup of the divergence angle of the transmitters. Therefore, a combination of both technologies has been shown to be advantageous over a homogeneous architecture also in normal operating conditions [140].

Other usage of hybrid RF/FSO systems, in particular referring to the normal operational state, include the adaptation of RF links for control purposes (to act as control channels) while utilizing FSO links for data transmission (due to their capacity being remarkably higher than that of RF links) [140,147].

There are several research papers focusing on the design of hybrid RF/FSO networks. The respective timeline of the evolution of the representative RF/FSO schemes is presented in Fig. 13. In particular, routing and traffic engineering is the main topic in Ref. [151], where a routing framework for hybrid RF/FSO networks is proposed. In particular, the authors introduce the concept of obscuration-tolerant paths (i.e., working paths protected by instantaneous backup paths) applied to serve the traffic in a weighted max-min way. Based on the evaluation presented in Ref. [151], an improvement of about 26% concerning the throughput can be achieved over the considered reference scheme used in commercial deployments.

In Ref. [152], a resilient routing scheme for hybrid RF/FSO networks based on the use of alternate paths is proposed. To provide the almost uninterrupted connectivity after a failure of an FSO link, the alternate paths are by default established also in the FSO domain. Switching the transmission to the RF domain takes place only if there is no possibility to restore the traffic via the alternate FSO paths.

The problem of topology control in hybrid RF/FSO networks is addressed in Ref. [153]. The introduced scheme of the RF/FSO network topology control utilizes the adaptive adjustments to the optical beamwidth of FSO transmitters at individual nodes and the regulation of the transmission power of the RF and FSO transmitters. An Integer Linear

<table>
<thead>
<tr>
<th>Vulnerability to:</th>
<th>FSO</th>
<th>RF</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) fog</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>(b) snow</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>(c) clouds</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>(d) rain</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>Line-of-sight conditions necessary</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Ability to diffract, reflect, penetrate, scatter on the obstacles</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>Risk of signal interference</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>License-free communications</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>Directional communications</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>Easiness of maintaining the connectivity with neighboring nodes</td>
<td>–</td>
<td>+</td>
</tr>
</tbody>
</table>

Table 3
Complementary characteristics of FSO and RF systems.
The routing performance under heavy rain and dense fog is addressed in Ref. [154], where the optimization model for resilient routing in hybrid RF/FSO networks is proposed. In particular, the authors present a class-based approach where routing of critical traffic demands is determined first using a pair of RF and FSO identical paths for each demand. Paths for non-critical demands are established in the second stage using the remaining network capacity (served only in the FSO domain by a pair of end-to-end link-disjoint paths). Due to this assumption, the critical traffic is offered the possibility to be protected against both rain and fog (due to the establishment of parallel paths in both FSO and RF domains). The non-critical traffic is similarly protected against rain (as the respective paths are established in the FSO domain) but its vulnerability to fog depends on the size of the area covered by the fog (in this scheme the non-critical traffic is impacted by fog, if both working and backup FSO paths are simultaneously impacted by the fog).

A similar problem of transmission in harsh weather conditions was investigated in Ref. [155], where a scheme to implement the parallel transmission via additional RF (or FSO) links was proposed. In that paper, the resilience of communications is enhanced by utilization of the parallel transmission paths (via the additional RF or FSO links) and by the use of the adaptive channel coding. Based on the analysis presented in Ref. [155], the best performance can be achieved for the case of a variable transmission rate of RF links to respond dynamically to the actual impact of weather conditions on the performance of the FSO links.

Hybrid RF/FSO architectures are also proposed in the literature for communications in space-air and satellite networks. For instance, routing for space-air networks consisting of satellites, aircrafts, and UAVs is addressed in Ref. [156], where a method of topology control to mitigate the impact of bad weather conditions is proposed. This method uses the strategy of the adaptive RF/FSO switching based on weather forecasts. It means that in the case of the predicted forthcoming bad weather conditions in a given area (e.g., incoming dense fog), FSO links are dynamically replaced by RF links. Routing is provided by means of the minimum cost paths based on the link cost function proposed in that paper to reflect the transmit power, the flow served by the link, link capacity, the initial and residual energy of nodes incident to a given link related to the transmit power, and the transmission time. According to the evaluation presented in Ref. [156], the resulting performance concerning the network lifetime and throughput can be rated high.

Similarly, in Ref. [157] a method of resilient routing for satellite constellations is proposed for a network with the inter-satellite communications performed via FSO links and communications between satellites and ground stations provided by either FSO or RF links. In this scheme, each FSO link is protected against the impact of the adverse weather conditions by the respective detour paths via other satellites using either the FSO or the RF links.

Finally, in Ref. [158], a multi-objective scheme is proposed to determine the topology and resource allocation together with the assignment of interfaces, the allocation of transmission links and channels, as well routing and topology control. In particular, in the context of weather events, the respective two-stage optimization problem is defined aimed at maximizing the network throughput by a gradual upgrade of the bottleneck RF links by FSO links. The authors show that by a proper FSO link augmentation, a significant improvement in terms of the network throughput can be achieved.

The multitude of deployment scenarios and the performance characteristics of free-space optical systems discussed in this section make FSO communications one of the key solutions for 5G (and beyond) wireless networks.

6. High-power jamming attacks

Attacks that target the availability of the network (resembling “denial of service” attacks), are the focus of this section, as these attacks have the potential to spread throughout the network, possibly impacting a large number of network connections. The main method for perpetrating such an attack is through the introduction of a high-power signal at the input port of a network switching node. This is called a high-power jamming attack, where a malicious optical signal is inserted at a network node, having an optical power much higher than that of the regular communication signals, and subsequently (through crosstalk at the network node) negatively affecting other signals (lightpaths) that also traverse the node that has been attacked. Essentially, this type of attack aims at exploiting the vulnerabilities of fiber-optic components within the optical network. Specifically, optical switches exhibit non-ideal characteristics of free-space optical systems discussed in this section make

Fig. 13. Timeline of the evolution of the representative RF/FSO schemes.

Programming model is presented with the objective to obtain the required QoS level. To be able to obtain the solution in the acceptable time for large-scale networks, Lagrangean relaxation is proposed along with iterative heuristics.

In this section, initially the modeling of the jamming attacks in optical networks is presented, followed by the description of a number of techniques on how to mitigate the impact of a jamming attack. These include the development of attack-aware routing and wavelength assignment (RWA)/routing and spectrum allocation (RSA) algorithms for WDM and EONs respectively, in the network planning phase, the placement of WSSs/spectrum selective switches (SSSs) at the input ports of the network switching nodes (as this would have been the case for opaque network architectures), essentially enabling the high-power jamming attacks to propagate to different parts of the network (as the signal does not undergo the 3R regeneration (i.e., re-amplification, re-shaping and retiming) at the optical switching nodes which would essentially put a stop to the spread of the attack).

In this section, initially the modeling of the jamming attacks in optical networks is presented, followed by the description of a number of techniques on how to mitigate the impact of a jamming attack. These include the development of attack-aware routing and wavelength assignment (RWA)/routing and spectrum allocation (RSA) algorithms for WDM and EONs respectively, in the network planning phase, the placement of WSSs/spectrum selective switches (SSSs) at the input ports of the network switching nodes (as this would have been the case for opaque network architectures), essentially enabling the high-power jamming attacks to propagate to different parts of the network (as the signal does not undergo the 3R regeneration (i.e., re-amplification, re-shaping and retiming) at the optical switching nodes which would essentially put a stop to the spread of the attack).
High-power jamming attacks in optical networks can be categorized as in-band jamming attacks (due to intra-channel crosstalks), out-of-band jamming attacks (due to inter-channel crosstalks), or gain competition attacks in optical amplifiers [159,160]. In the first two cases the interactions between connections at common nodes or links through either in-band or out-of-band channel crosstalk can be used to model the propagation of high-power jamming attacks at the optical layer. In the third case, a high-power jamming signal is introduced at an optical amplifier, which as a result will force the rest of the signals on the same fiber (at the input of the optical amplifier) to experience lower gain. Given the fact that transparent networks are considered, this effect will become more pronounced as the signals pass through consecutive optical amplifiers (increasing the power of the malicious signal that experiences higher and higher gain, and lowering the power of the rest of the co-propagating signals that experience lower and lower gain).

An in-band jamming attack can occur at a network node if the malicious signal is interfering (due to intra-channel crosstalk) with another communication signal at the same wavelength (or the central nominal frequency for EONs) – Fig. 14(a). For the case of an out-of-band jamming attack, the high-power signal that is injected at the input port of a switching node can introduce nonlinearities in the network, subsequently causing crosstalk effects between channels on different (adjacent) wavelengths that are co-propagating with the malicious signal on the same fiber (inter-channel crosstalk) (Fig. 14(c)). Between these two types of attacks, it is evident that in-band jamming is the most severe attack, as it is not possible to use a filter to remove the intra-channel crosstalk (both attacking and attacked signals are on the same wavelength).

On the other hand, experimental results have shown that inter-channel crosstalk becomes more severe when different line rates and modulation formats are utilized in the network (a common practice in today’s networks). Thus, both intra- and inter-channel crosstalk must be carefully considered and mitigated during the network planning phase, to ensure that any high-power jamming attack will not affect the network performance. A third type of attack, as previously mentioned is gain competition (Fig. 14(b)). Such an attack must be also mitigated as it will have a significant impact on the signal especially when it traverses several amplifiers in series (over a long fiber span).

As previously mentioned, jamming attacks can have a disastrous impact on the network operation, as they can potentially spread throughout the network. This is possible, because the attacked lightpath now becomes a secondary attacker, affecting any other lightpath it overlaps with (utilizing the same or adjacent wavelengths) at other network nodes or links; these lightpaths subsequently affect other lightpaths, and so on. Clearly, jamming attacks can affect a large number of lightpaths, causing disruption of service in large segments of the network.

The reader should note that, without loss of generality, for the rest of the section, only WDM networks will be considered (i.e., only the terminology for WDM networks will be utilized henceforth in this section). Nevertheless, the same techniques/approaches described below apply to EONs, too.

For modeling the in-band and out-of-band jamming attacks, initially the topology of the network is modeled as an undirected graph \( G = (V, E) \), where the set of vertices \( V \) represents the set of optical nodes and the set of edges \( E \) represents the set of fiber links. Each fiber link supports \( W \) distinct wavelengths. Subsequently the propagation of the attacks is modeled as the interaction between different lightpaths at the network nodes and/or the network links.

Fig. 15 illustrates an example of a high-power in-band jamming attack at node \( n_1 \) using wavelength \( w_1 \). This malicious signal will initially affect lightpath \( p_0 \) that also utilizes wavelength \( w_1 \). Lightpath \( p_0 \) will subsequently become a “secondary attacker”, further spreading the attack to lightpath \( p_2 \) (also on \( w_1 \)).

Similarly, Fig. 16 illustrates how a high-power jamming attack can propagate via the inter-channel crosstalk. In this case, node \( n_1 \) is attacked using wavelength \( w_{i+1} \). This malicious signal will affect lightpath \( p_0 \) (on wavelength \( w_i \), since they use adjacent wavelengths on the same fiber link). Similarly to the in-band jamming attack example, lightpath \( p_0 \) becomes a “secondary attacker”, consequently affecting lightpath \( p_2 \). The reader should note that in this example only adjacent wavelengths can be attacked (thus, lightpath \( p_2 \) is not affected).
6.2. Selected techniques of jamming attack mitigation

As described above, crosstalk interactions among different connections in the network enable the spreading of malicious high-power signals in the network. Thus, techniques must be developed in terms of network planning, such that the impact of a jamming attack on lightpaths other than the one attacked, is minimized. There are two main approaches in order to mitigate the impact of the jamming attacks and ensure service availability against any potential service disruption for the network connections. The first approach is to utilize the attack-aware routing and wavelength assignment optimization algorithms and/or heuristics during the network planning phase to provision the connections in such a way that minimizes the impact of any jamming attack that may take place in the network (i.e., minimizes the interaction among lightpaths and as a consequence reduces the spread of the attack). Another approach is to use wavelength-selective attenuators as power equalizers inside network nodes or place WSSs at specific network switching nodes in order to limit the propagation of the introduced high-power jamming signals. Finally, OPM equipment can be placed at specific network locations, that can monitor the crosstalk interactions and provide attack-awareness. This information can be subsequently utilized to mitigate the effect of an intentional attack.

6.2.1. Attack-aware connection provisioning

The most cost-effective (in terms of capital expenditures – CAPEX) technique to address the high-power jamming attack problem is to provision the connections (i.e., solve the routing and wavelength assignment problem) such that in the event of a jamming attack the number of interactions between lightpaths (either due to inter- or intra-channel crosstalk) will be kept to a minimum. This way, the attack will not propagate throughout the network, ensuring the availability of the network connections. Obviously, when such a technique is employed, care must be taken to ensure that the paths and wavelength assignments found are such that they do not negatively affect the performance of the network, i.e., the paths found must not be long and the resources must be allocated efficiently. The former will ensure that the latency experienced within the network is kept low, and the latter will guarantee that the blocking rate will also be kept low (as blocking due to resource exhaust will be experienced for higher network loads).

There have been several works in the literature addressing this problem. Initial work appeared in Refs. [159,162], followed by works in Refs. [163,164]. In general, in Refs. [159,162], the authors presented the idea of preventing the attack by appropriately solving the RWA problem. For example, the work presented in Ref. [159] focused on the development of an ILP formulation for the routing problem, aiming to find paths such that the potential out-of-band crosstalk and gain competition at the optical amplifiers is minimized, while a metaheuristic (tabu search) was developed to address larger networks that could not be solved by the ILP. This work was extended in Ref. [163] to also include in-band jamming attacks, and it was further extended in Ref. [164] where the propagation of the attack by the secondary attackers was limited due to practical considerations.

Extending the aforementioned approaches, the authors in Refs. [161, 165] proposed an optimization approach that jointly solved the RWA problem, while also considering for both in-band and out-of-band crosstalk interactions. The final aim was to minimize the number of interactions, thus limiting the spread of the attack. In that work, linear programming (LP) relaxation techniques, as well as heuristic algorithms were proposed, in order to address larger size problems. A meta-heuristics was again used (simulated annealing in that case) that found the best ordering for the connections that were sequentially serviced by the heuristic algorithm. It is interesting to note that the proposed techniques achieved to limit significantly the spread of the attack, while requiring only a slightly higher number of resources compared to the non-attack-aware case (that aimed to maximize the resource efficiency). These results clearly validate the proposed techniques and demonstrate that if the appropriate allocations take place during the planning of the network, then any single jamming attack can be treated in an efficient and effective manner.

Finally, as also mentioned in Sections 6.2.2 and 6.4 that follow, the work presented in Ref. [166] examined jointly also the problem of WSS placement while the work presented in Ref. [167] addressed the joint problem of minimizing the spectrum utilization, the number of required WSSs, and the number of required lightpath re-allocations in the presence of traffic uncertainties. Different results can be obtained for the latter optimization, based on the choice of certain coefficients (weights) used in the objective function, that depend on the network operator’s requirements in terms of operating expenditures (OPEX) and CAPEX. Nevertheless, it is expected that a solution that keeps low the number of WSSs, the number of re-allocations, and the bandwidth utilization, will be preferred.

6.2.2. Equalizer/WSS/OPM placement

The placement of equalizers, WSSs, or OPMs at specific network locations in conjunction with the implementation of attack-aware optimization algorithms and heuristics is another technique to mitigate the propagation of jamming attacks.

Initial work that appeared in Ref. [168] and then subsequently in Ref. [169], investigated this problem. Specifically, wavelength-selective attenuators were used inside network nodes as power equalizers in order to equalize the signals and mitigate the propagation of the high-power signals. These wavelength-selective attenuator modules, placed before the optical amplifiers located at the output of the nodes, are comprised of photodetectors, variable optical attenuators, and a control entity to dynamically control the signals’ power levels. Clearly, the use of such a module allows the network operator to eliminate the problem of gain competition at the optical amplifiers, as well as avoids the problem of inter-channel crosstalk along the fiber links. As also discussed in Ref. [168], optical limiting amplifiers can also be used that are able to provide constant output power over a wide range of input power variations, thus mitigating the gain competition and inter-channel crosstalk effects. However, such a solution may be very costly. Also, it will not be sufficient in the case where the input power variations between the attacking and attacked signals exceed the range of the limiting amplifier.

The goal of both aforementioned works was the minimization of the number of equalizers, thus minimizing the network’s CAPEX. Initially, in Ref. [168], greedy algorithms were developed that aimed at reducing the propagation of the attack, while limiting the number of equalizers that were placed sparsely within the network. In addition, extension of that work, presented in Ref. [169], provided the optimal solution (via an ILP formulation) that had as an objective the minimization of the number of equalizers placed in the network. Such a solution is clearly not applicable for larger-size networks, where the greedy approach can not be employed.

The equalizer placement problem to mitigate jamming attacks was also investigated in Ref. [170], where a different node architecture was introduced (based on WSSs) achieving equalization utilizing optical attenuators with a feedback loop. Another key feature of the proposed architecture was its modularity, allowing the equalizers to be placed at specified modules in the node (avoiding assigning equalizers for the entire node as in previous approaches). Thus, these equalizers were now not only allocated to specific nodes but also to specific modules within
the node. In this work, the joint connection provisioning and equalizer placement problem was also addressed via an ILP formulation (contrary to previous approaches that addressed these two problems independently), in an effort to minimize both the propagation of the attacks (i.e., minimize both intra- and inter-channel crosstalk), as well as the required number of equalizers.

Another important consideration in trying to minimize the propagation of the high-power jamming attacks is the port isolation in optical switching nodes (e.g., ROADMs), especially in today’s networks where the size of the switching nodes increases, thus, enhancing the crosstalk effect. An easy solution would be clearly to replace the network switching nodes with ones that have a very high port isolation; such a solution though would be prohibitively expensive. An alternative cost-effective solution is to upgrade only a few (specific) switching nodes (ROADMs) by replacing some of the splitters at the input stages of the ROADMs (broadcast-and-select (BS)-based ROADM architectures are assumed) with WSSs (creating what is known as “Architectures on Demand” (AoD) – see Fig. 17).

Similar to the work presented for the equalizer placement problem, in order to reduce cost but at the same time limit the attack propagation, the appropriate provisioning algorithms in conjunction with the placement of WSSs have been considered. In this case, the WSSs have lower port isolation characteristics, and thus consequently have a lower cost compared to the high-port isolation solution outlined above. Nevertheless, it should be noted that signal suppression is possible utilizing this architecture, as there exist two points of suppression for all paths that experience crosstalk, ensuring that the required isolation is achieved.

Specifically, in Refs. [166,167], ILP formulations were developed to minimize both the number of crosstalk interactions within the network through the appropriate provisioning of all network connections at the network planning phase (both intra- and inter-channel crosstalk interactions were considered), as well as to minimize the number of WSSs to be placed within BS-based ROADMs (by replacing specific splitter units at the input stage of the ROADMs with WSSs). A heuristic technique based on vertex coloring was also developed in Ref. [166] to address the scalability issues related to the ILP. It was shown that the proposed algorithms were able to mitigate the attack propagation, while utilizing the minimum number of WSSs, with only a slight penalty in terms of spectrum utilization.

Finally, in a similar vein, work in Ref. [171] jointly addressed the problem of attack-aware connection provisioning and sparse placement of OPM equipment in WDM networks via an ILP formulation and a genetic algorithm (GA). It should be noted that information provided by the OPMs, that can efficiently detect an attack as they are monitoring signal quality, can be indirectly used to minimize the impact of an attack (e.g., by taking specific actions to enhance signal quality for a connection under attack). The scalable metaheuristic (GA) presented minimized crosstalk interactions (with results close to the optimal) and utilized a small number of wavelengths and OPMs, thus keeping the network cost low. This is achieved by selecting the fittest solution in the population (i.e., the smallest cost chromosome), utilizing a fitness function that takes into account the number of required wavelengths, the number of in-band and out-of-band crosstalk interactions, as well as the spread of the interactions across the output ports of the affected network nodes that is used to minimize the number of required monitors.

6.3. A critical comparison

Depending on which technique is used (i.e., attack-aware provisioning, attack-aware provisioning in conjunction with equalizer placement, and attack-aware provisioning in conjunction with WSS placement) as well as the method in which the approach is implemented (i.e., ILP, LP relaxation, heuristic algorithms, heuristic algorithms with metaheuristics (e.g., tabu search, simulated annealing)) a number of tradeoffs can be observed. For example, ILPs will provide optimal solutions (i.e., in terms of the resource utilization, number of required components, number of lightpath interactions) but are slow and cannot scale to large-sized networks, while heuristics or LP relaxations provide fast (sub-optimal) solutions and can easily scale to large networks. When metaheuristics are also utilized (to sort for example the list of connection requests) solutions closer to the optimal can be obtained, at the expense of the additional computational cost.

Further, if attack-aware provisioning is used without placing additional network components (e.g., equalizers, WSSs), the solutions obtained in terms of the number of lightpath interactions (i.e., the attack propagation) will be worse than in the case when additional network components are placed at specific locations in the network (e.g., WSSs replacing splitters in BS-based ROADMs or equalizers (implemented as optical attenuators with a feedback loop) placed in specific node modules). In the latter case, these additional components serve to compensate for the crosstalk effect. Clearly though these solutions entail additional CAPEX and in the case of equalizer placement dynamic control of the power levels is also required.

In general, the attack mitigation technique as well as the implementation method that will be utilized for this technique, constitute interesting questions, the answers of which depend on the network operator’s objectives (e.g., best meeting specific network performance targets), the end-users’ needs, and the traffic demand behavior (in the case of traffic demand uncertainty as well).

6.4. Future directions for jamming attack mitigation

In most of the works presented in the literature it is assumed that there is no uncertainty associated with the traffic demand and connections tend to stay in the network for a long period of time. However, this is not usually the case in real networks, where traffic demands vary with time. For example, traffic demand patterns may exhibit behavior similar to the Internet traffic demand that is usually modeled by the log-normal distribution [167,172–174]. In this case, the jamming attack mitigation techniques utilized (either attack-aware provisioning techniques or equalizer/WSS placement techniques) must now also consider the time-varying traffic demands. Further, today’s networks such as the IoT and 5G networks, as well as new applications and services that are based on such networks, introduce a tidal traffic effect to the network traffic [175,176]. This is the case, since now traffic fluctuations occur due to the movement of people to different areas at different times. Thus, this tidal traffic effect must again be modeled appropriately, and subsequently security against jamming attacks must be provided now to
SDN-enabled reconfigurable optical networks that can adapt to these traffic demands [176-178].

Another future research avenue includes the consideration of jamming attack mitigation techniques for more recent network architectures, such as spectrally-spatially flexible optical networks (SSFON) utilizing a number of different optical fiber schemes (such as multimode fibers (MMFs), few mode fibers (FMFs), multiple core fibers (MCFs), or bundles of single core fibers), as well as a variety of switching mechanisms (such as independent switching, joint switching, and fractional joint switching) [179]. Clearly, based on the fiber and switching technology used, the impact of a jamming attack will differ. A recent work addressed the impact of high-power jamming attacks on space division multiplexing (SDM) networks utilizing MCFs [180], demonstrating that inter-core crosstalk enables the spreading of the attack to multiple cores. The exact impact of the crosstalk will depend on specific physical characteristics of the fiber such as the distance between adjacent cores, whether adjacent cores operate at the same nominal central frequency, and the type of modulation format used at each frequency slot (for EONs). Thus, any jamming attack mitigation technique proposed will offer benefits and drawbacks regarding OPEX (e.g., performance and flexibility), as well as CAPEX (e.g., equipment costs).

In addition, as failure protection and security from malicious attacks are not mutually exclusive, the goal should be to develop techniques that jointly solve both problems. Recent works in the literature have tackled this joint problem for the case of eavesdropping attacks (where network coding is utilized to provide security, with backup paths found to protect the network in the event of any single link failure, while preserving the level of security provided) [64], as well as jamming attacks [181, 182]. Specifically, in Ref. [181] a jamming attack-aware dedicated path protection algorithm is developed that establishes disjoint working and backup paths for each connection, while in Ref. [182] that work was extended to provide a more spectrum-efficient solution by developing a jamming-aware shared path protection optimization algorithm that provides both security from jamming attacks, as well as protection against any single link failure without requiring additional resources (in terms of bandwidth usage) compared to the case when only protection is provided. Clearly, this is a complex problem that requires attention from the research community, as in addition to attack-aware provisioning of the primary path for each connection request, backup paths (and their corresponding spectrum assignment) must now also be obtained again aiming to minimize the spread of a potential attack, while also minimizing network resources and additional equipment costs. Further, security against jamming attacks in the presence of multiple failures is another problem that needs to be considered.

Finally, in dynamic optical networks, enabled by SDN, statistical models, as well as machine learning methods can be applied (e.g., neural networks, deep neural networks, and reinforcement learning). Initial works in this area have recently appeared in the literature focusing mainly on the detection and identification of jamming attacks [183], with a more current work addressing the jamming attack prevention problem as well [184]. Specifically, performance results in Ref. [184] demonstrated that the usage of artificial neural networks, amongst different machine learning techniques, was the fastest and more accurate method for detecting a jamming attack and identifying which optical channel was attacked. This was also the case in Ref. [183] where machine learning approaches were investigated utilizing the experimental attack traces obtained from a field testbed, once again demonstrating that artificial neural networks achieve the best performance in terms of accuracy in identifying the type of attack. In terms of prevention, a resource reallocation scheme was also proposed in Ref. [184] that utilizes statistical information of detection accuracy to achieve a lower jamming probability, while also minimizing the number of lightpath re-allocations (which, in turn, minimizes the service disruption in the network). Clearly, this is a vast research area that warrants a significant effort in order to obtain solutions that, utilizing an appropriate machine-learning framework, can provide fast and accurate predictions on attack detection and identification, as well as attack mitigation strategies that are both effective and efficient in terms of network resources and network operation.

7. Strategies for post-disaster recovery

Earlier works in post-disaster recovery (PDR) span survivability-driven approaches that aim for bandwidth recovery and fast re-provisioning whereas recent studies focus on fixed-mobile convergence/integration and data science-driven strategies with the advent of 5G communications and data-enabled computational solutions. Existing methodologies for post-disaster recovery in optical networks include optimization-based solutions such as ILP and/or mixed integer linear programming (MILP) models, dynamic programming-based solutions, heuristics and meta-heuristics. This section begins with a brief introduction of modeling and assessment of disasters for post-disaster recovery solutions, and continues by overviewing the existing studies in post-disaster recovery of optical networks, a comparison of the existing solutions, and a thorough discussion on the way forward, focusing on yet open issues, challenges and opportunities in this field. The organization, as well as a high level overview, of this section is illustrated in Fig. 18.

7.1. Modeling of post-disaster scenarios

To model the post-disaster recovery scenario, the impact of the disaster needs to be defined properly. To help design communication networks in response to disasters, the study in Ref. [185] categorizes disasters into human-made and natural disasters, whereas the latter includes geophysical, hydrological, climate-related, weather-related and biological disaster sub-categories. Although it is hard to model the post-natural disasters, as suggested by the study, human-made disasters are more difficult to predict; hence, post-disaster models of natural disasters are vital for communications networks. To this end, the study in Ref. [186] introduces the degree of network damage (DND) metric. DND quantifies the impact of the disaster on the network infrastructure, humans, and traffic flows. The discrete quantification of the DND corresponds to one of five degrees. Out of these five degrees, DND Level 1 denotes no network outage whereas DND Level 5 denotes 100% network outage probability with a repair time varying from 14 to 25 h.

Disaster information management is of paramount importance. To this end, the study in Ref. [187] presents a requirement analysis of disaster management systems alongside the potential data-driven methodologies to boost the situation awareness in post-disaster recovery, while meeting the user needs at the same time.

In Ref. [188], the network and communication restoration is proposed to be managed in a context-aware and delay tolerant manner. To this end, post-disaster situation modeling mostly relies on pre-defined categories of trackable movement patterns of different types of users that require connectivity services. It is proposed that a traffic engineering solution should take into consideration the following mobility models: Post-Disaster mobility model [189] which comprises four mobility models, particularly recurrent motion for transportation, localized random motion of rescue teams, recurrent motion of security patrols, and a motion switching between a center and random destinations for paramedic services. In addition, the Valparaiso mobility model [190] is considered which models security points, evacuation routes, and emergency mobility patterns published by some governments.

7.2. Remarkable schemes for post-disaster recovery

Table 4 presents remarkable schemes for post-disaster recovery.

Post-disaster recovery strategies in these works aim for virtual network (VN) restoration, best-effort fairness, reliability-bandwidth trade-off, differentiated restoration, cloud network mapping, datacenter restoration, datacenter reliability, virtual network service recovery, carrier
restoration, data evacuation, fixed-mobile cooperation, and open and disaggregated subsystems. Methodologies include Integer Linear Programming (ILP)/Mixed ILP (MILP)-based optimization, heuristics, traffic engineering and binary integer optimization. While most schemes do not specify the disaster types, specific disaster types in a few works include earthquakes, tsunamis, human-made destructions and power grid failures. Below, we provide a summary of these reviewed schemes in detail.

As a survivability-based study, the authors in Ref. [191] tackle the post-disaster recovery of optical VNs by quantifying the damage cost in terms of disconnected VNs, failed virtual links, and failed physical links. To this end, an MILP-based optimization model is proposed as a baseline, as well as heuristics that include dynamic programming and simulated annealing techniques.

In disaster preparedness, reliability and throughput yield a trade-off. To cope with this, the study in Ref. [192] proposes a disaster preparedness strategy that leverages the reliable route discovery methodology in optical networks with the objective of bandwidth recovery for the connections affected by a disaster. To this end, a reliable bandwidth recovery heuristic is proposed to achieve a trade-off between bandwidth utilization and service reliability in the case of a natural disaster.

In an earlier study [193], the authors propose a best-effort fairness degradation based multipath re-provisioning heuristic which follows an intuitive strategy by assigning minimum bandwidth to each connection followed by cyclical upgrades to bandwidth assignments. As an alternative, in the same study, the authors propose a best-effort fairness

<table>
<thead>
<tr>
<th>Table 4</th>
<th>Remarkable post-disaster recovery (PDR) schemes in optical networks.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post-Disaster Recovery Strategy</td>
<td>Basis</td>
</tr>
<tr>
<td>Virtual network restoration [191]</td>
<td>Survivability</td>
</tr>
<tr>
<td>Best-effort fairness [193]</td>
<td>Multi-path survivability</td>
</tr>
<tr>
<td>Reliability-bandwidth trade-off [192]</td>
<td>Reliable route</td>
</tr>
<tr>
<td>Multi-path restoration + bitrate squeezing [203]</td>
<td>Restorability</td>
</tr>
<tr>
<td>Differentiated restoration [204]</td>
<td>Restorability</td>
</tr>
<tr>
<td>Cloud network mapping [202]</td>
<td>Survivable cloud network mapping</td>
</tr>
<tr>
<td>Datacenter restoration [198]</td>
<td>Max. connectivity</td>
</tr>
<tr>
<td>Datacenter reliability [199]</td>
<td>Min. resources</td>
</tr>
<tr>
<td>VN service recovery [200]</td>
<td>Min. recovery latency, max recovery</td>
</tr>
<tr>
<td>VN restoration [201]</td>
<td>Max. VN mapping and recovery</td>
</tr>
<tr>
<td>Carrier restoration [205]</td>
<td>Multi-carrier collaboration</td>
</tr>
<tr>
<td>Data evacuation [206]</td>
<td>Software-defined satellite networks</td>
</tr>
<tr>
<td>Fixed-mobile cooperation [208]</td>
<td>Service restoration</td>
</tr>
<tr>
<td>Open and disaggregated subsystems [209,210]</td>
<td>Recreation of the lost performance monitoring and perform robust telemetry</td>
</tr>
</tbody>
</table>
degradation based multipath route search for improved bandwidth allocation during a post-disaster scenario. With the advent and widespread use of cloud-based services, datacenters (DCs) have become the hosts of cloud systems. Optical networks have been envisioned to ensure inter-datacenter connectivity as reported by several researchers [194,195]. Besides the disaster preparedness solutions for optical inter-datacenter networks [196,197], the state of the art also offers solutions for post-disaster recovery of the optical backbone interconnecting DCs. The study in Ref. [198] proposes a heuristic solution to improve network connectivity while maximizing the efficiency of DC restoration in the case of a disaster affecting one or more zones where DCs are located. Similarly, the study in Ref. [199] tackles the problem of post-disaster recovery of an inter-datacenter optical network. An ILP formulation is proposed to maintain the requests for DCs by utilizing minimum network resources after a disaster. Apart from the other studies in the literature, the authors investigate the impact of the number of DCs, optical reach, and the category of disasters for DCs by utilizing minimum network resources after a disaster.

To this end, a heuristic and a meta-heuristic are used to determine the near-optimal placement of repair resources to restore VN services in an incremental manner. Recovery rate and speed of VN services are considered as the key performance indicators in that study. The intuition behind formulating the post-disaster recovery as a network restoration problem is that a disaster which leads the substrate network to fail will propagate the failures towards the VN demands. While the study in Ref. [200] also proposes a virtual link and node recovery strategy based on an ILP optimization model for an inter-datacenter optical WDM network, it differs from other similar studies by considering cascading failures following upon a disaster, and building their solution on a failure propagation model of the power grid network. The performance of the proposed solution has been shown to be efficient in terms of the number of mapped/recovered and disconnected cloud networks following the cascading power grid failures. As most studies either do not consider or focus on a single disaster scenario, the work presented in Ref. [201] also considers an earthquake scenario and a human-made massive destruction disaster scenario separately for the post-disaster recovery of a cloud network realized on opaque WDM links. By introducing an ILP formulation, the presented work aims at two objectives: survivable post-disaster embedding of a cloud network and post-disaster economic sustainability for the operators, with the cloud network embedding problem building on risk assessment and VM backup placement.

Considering post-disaster conditions, the study in Ref. [202] proposes a heuristic algorithm for single path restoration of an SDN-based EON backbone with the maximum restorability objective. By extending this idea for an EON in a post-disaster setting, the study in Ref. [203] proposes a multi-path heuristic that leverages the differentiated class of service concept by taking into account the delay-tolerance constraint for post-disaster restoration and maximum tolerable bandwidth degradation per optical connection. The differentiated restoration introduced by the authors aims at coping with multiple failures in a post-disaster situation.

Post-disaster recovery can benefit from SDN in many ways. As reported in Ref. [204], carrier collaboration is one of the remarkable benefits that can facilitate post-disaster recovery. With this in mind, the authors of that study proposed an SDN-based architecture to enable multi-carrier collaboration so as to form an emergency packet transport network. This is referred to as “emergency exchange points of carrier collaboration (EPOC)” in the study. Through an experimental testbed, it was shown that the EPOC-based post-disaster recovery solution can guarantee meeting the traffic requirements, while ensuring fast recovery for the carriers at lower costs. Another benefit of SDN is presented as the potential to define an SDN controller to evacuate data from DCs following a disaster [205]. Thus, the authors of the study approach the post-disaster recovery problem not only from the network connectivity standpoint but also from a data-centric point. To this end, software defined satellite networks are utilized to evacuate data from DCs towards their next safe destinations in the backbone network. Through a traffic engineering approach that builds on a graph-theoretic solution that generates an event-driven graph from a contact graph, the proposed SDN-based satellite networking solution is shown to outperform a delay-tolerant networking-based approach in terms of evacuated data within the first half-day following a disaster.

As mentioned earlier, fiber and wireless integration is inevitable in the 5G era [206]. With this in mind, the study in Ref. [207] builds on network cooperation protocols so as to deploy a recovery node to bypass the disaster-impacted traffic through a wireless backhaul based on the disaster risks of potential optical nodes. While no particular disaster is considered, a conceptual strategy is introduced via binary integer optimization methodology to select the recovery nodes in the optical network in the case of a disaster. The objective of post-disaster recovery is set as the maximum throughput over wireless links constrained to the disaster risk of potential recovery nodes in the network.

7.3. Open issues, opportunities and future directions

Despite the ongoing efforts and existing solutions for post-disaster recovery of optical networks, challenges and open issues remain for various aspects. Below, we provide useful insights for the researchers who aim to tackle challenges in this field.

7.3.1. Realistic models for a post-disaster context

Existing studies that were reviewed earlier mainly focus on network failure and survivable restoration solutions for the networks. On the other hand, the mobility models concerning the post-disaster scenarios (see Section 7.1) are expected to transform the traffic demand significantly. That being said, the state of the art calls for realistic models of evacuation and/or temporary movements that would translate to dramatic changes in the traffic demand patterns. The differentiated restoration concept in Ref. [208] can significantly benefit from a realistic model of user mobility and varying traffic demands since the services provisioned through the first responders’ connectivity are considered to be of the highest priority whereas operators must commit to a certain degree of restorability of user services based upon their contracts.

7.3.2. Leveraging fixed-mobile convergence in 5G for robust data evacuation

The study in Ref. [209] suggests the software-defined satellite communications to evacuate data during the post-disaster recovery phase. Despite the study being unique for considering data evacuation in post-disaster recovery, it is worth noting that aerial communications under the 5G ecosystem can ensure wide radio coverage to help fixed (e.g., wireline, optical) and mobile terrestrial networks restore and/or evacuate data via aerial and space networks [210]. As reported by several researchers [211], fixed-mobile integration, more specifically emerging optical network solutions including 400+ Gb/s coherent modulation and detection, optical switches and next-generation ROADMs can ensure high-throughput and low-latency for the fronthaul and backhaul of the wireless 5G network. Fog-based radio over optical fiber networks are examples of low-cost optical-wireless integrated solutions in the 5G ecosystem [212]. While ensuring the restoration of an integrated fixed (optical)-mobile (radio) network following a disaster, the potential of UAVs needs to be investigated in post-disaster scenarios for data evacuation from DCs [213]. This requires the optimal deployment of aerial nodes in a post-disaster recovery scenario. Thus, when the optical inter-datacenter network is hit by a natural or human-made disaster, a real time or near-real time response to potential data loss in DCs should be addressed thoroughly by the optimization models or heuristics.
7.3.3. Data-driven adaptive restoration

As presented in Table 4, the majority of the studies do not differentiate solutions between the types of disasters. Thus, it is implicitly assumed that the outcomes of all disasters are identical from the network point of view. Although this might be an agreeable assumption for a feasibility or proof-of-concept study, it is worth noting that each type of a disaster reveals its own failure signature on the communication infrastructure and calls for a particular decision support [215]. With this in mind, to ensure a low cost and highly efficient recovery of optical networks following a disaster, context-aware (disaster-specific) restoration schemes are emergent. Indeed, context-awareness involves building and training machine learning models to recognize the post-disaster situation precisely. Thus, post-disaster recovery of optical networks calls for predictive analytics-backed decision support to adapt to the post-disaster situation (i.e., context) based upon the pre-determined restoration schemes.

7.3.4. Computational efficiency in post-disaster recovery

The majority of the solutions rely on optimization models, more specifically ILPs and/or MILPs. Although these formulations can provide credible benchmarks for the post-disaster recovery strategies, the state of the art calls for real-time and near-optimal solutions. That being said, heuristics that perform at a certain degree of optimality need to be improved by novel solutions that address the trade-off between computational efficiency and optimality in terms of the restoration/recovery objectives such as disaster damage cost, recovery time, traffic loss, fairness, connectivity, and throughput. Furthermore, it is worth to note that the state of the art also calls for holistic strategies that can imitate the optimization-based benchmark solutions without requiring high computing power.

8. Conclusions

In this paper, we focused on the presentation of a set of representative techniques to enhance the disaster resilience of optical networks concerning all major groups of external disaster events including natural disasters, weather-induced massive disruptions and malicious human activities. Our special focus was on presenting, apart from the proactive mechanisms applied before the occurrence of a disaster, also the mechanisms relevant for foreseeable disasters which can be used in a short period of time before the occurrence of disaster symptoms, and the time the disaster hits the network (such as, e.g., those involving data evacuation), as well as mechanisms of post-disaster recovery.

Our main conclusions are as follows. In the case of foreseeable disasters, it is possible to take advantage of the time between alerts are issued and the time of the actual occurrence of the disasters to undertake certain actions to mitigate the disaster impact. These actions include the migration of data (especially critical data) and VMs to safer locations. In this context, different pre-planned evacuation strategies may be considered.

A new dimension of disaster resilience refers to data evacuation operations possible, e.g., in the case of noticing the symptoms of an earthquake several tens of seconds before it hits the network. Concerning the optical transmission in the wireless domain using the OWS components, for resilience purposes, it is crucial to apply the mechanisms of resilience to adverse weather conditions, in particular including dense fog, snow, and clouds. For this purpose, an appropriate approach seems to be to use a hybrid RF/TSO architecture, as its advantages have been confirmed in many research papers.

In the context of resilience of optical networks to malicious human activities, in this paper we showed that high-power jamming attacks have the potential to create effects similar to denial of service attacks in large network segments if not mitigated properly. Therefore, it is imperative that such mitigation techniques are applied during the network planning phase to ensure such attacks do not spread unchecked if and when they happen. Both attack-aware provisioning techniques as well as the addition of extra network components (equalizers, WSSs) can be utilized to address such attacks. The choice ultimately depends on the network operator’s OPEX and CAPEX requirements.

Finally, the post-disaster modeling has been shown to be another key component in the design of recovery solutions. Among the post-disaster recovery strategies, cloud network mapping and DC restoration are vital, and they need to be coupled with effective data evacuation schemes, which require leveraging the fixed-mobile convergence within the SG and beyond era.

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

The work of R. Girão-Silva and T. Gomes was partially supported by Fundação para a Ciência e a Tecnologia (FCT), I.P. under project grant UIDB/00308/2020 and was financially supported by ERDF Funds through the Centre’s Regional Operational Program and by National Funds through FCT under project CENTRO-01-0145-FEDER-029312. The work of G. Ellinas was partially supported by the European Union’s Horizon 2020 research and innovation programme under grant agreement No 739551 (KIOS CoE) and from the Government of the Republic of Cyprus through the Directorate General for European Programmes, Coordination and Development. It was also partially supported by the Cyprus Research and Innovation Foundation under project CULTURE/award-YR/01418/0014 (REALFON). The work of B. Kantarci was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) DISCOVERY Program under RGPIN/2017–04032. Massimo Tornatore acknowledges the support by U.S.–Japan JUNO2 project (NSF Grant no. 1818972).

This article is based on work from COST Action CA15127 (“Resilient communication services protecting end-user applications from disaster-based failures” – RECODIS), supported by COST (European Cooperation in Science and Technology); http://www.cost.eu.

References


[98] M. Shahnavaz, V.P. Maheshwari, Global disaster data research and threat detection system, in: International Conference on Computational Intelligence and Knowledge Economy, Smart Power and Efficiency, Hamburg, Germany, pp. 1–6.


[102] M. Shahnavaz, V.P. Maheshwari, Global disaster data research and threat detection system, in: International Conference on Computational Intelligence and Knowledge Economy, Smart Power and Efficiency, Hamburg, Germany, pp. 1–6.


Teresa Gomes is an Assistant Professor at the University of Coimbra, at the Department of Electrical and Computer Engineering, since 1998, and with the permanent position since 2003. In 2013, from April until July, she was a Visiting Researcher at the School of Information Sciences of the University of Pittsburgh (USA). She is also a researcher at the Institute for Systems Engineering and Computers at Coimbra (INESC Coimbra). She received her PhD diploma in Electrical Engineering (Telecommunications and Electronics) at the University of Coimbra in 1998. She is the author/co-author of over 90 technical publications in international journals, conference proceedings, and book chapters, including one European patent. She has been responsible for several national projects, mainly with industry. She has also served as a TPC member of numerous international conferences and was Co-Chair of DRCN 2019. Teresa Gomes is an Associate Editor of Springer’s Journal of Network and Systems Management. Her main present interests are routing, protection and reliability analysis models and algorithms for communication networks.

Georgios Ellinas holds a B.S., M.Sc., M.Phil., and a Ph.D. in Electrical Engineering from Columbia University. He is a Professor and the Chair of the Department of Electrical and Computer Engineering and a founding member of the KIOS Research and Innovation Center of Excellence at the University of Cyprus. Previously, he served as Associate Professor of Electrical Engineering at City College of New York (2002–2005), Senior Network Architect at Tellium Inc. (2000–2002), and Research Scientist/Senior Research Scientist at Bell Communications Research (Bellcore) (1993–2000). Prof. Ellinas is a Fellow of the IET (2019), a Senior Member of IEEE and OSA, and a Member of ACM, and the Marie Curie Fellows Association (MCFA). He has co-authored/co-edited four books on optical networks, more than 245 archived articles, conference papers, and book chapters, and he is the holder of 30 patents on optical networking. His research interests are in optical/telecommunication networks, transportation networks, IoT, and critical infrastructure systems.

Burak Kantarci is an Associate Professor with the School of Electrical Engineering and Computer Science at the University of Ottawa. From 2014 to 2016, he was an Assistant Professor at the ECE Department at Clarkson University, where he currently holds a courtesy appointment. Dr. Kantarci received the M.Sc. and Ph.D. degrees in computer engineering from Istanbul Technical University, in 2005 and 2009, respectively. He received the Siemens Excellence Award in 2005 for his studies in optical burst switching. During his Ph.D. study, he studied as a Visiting Scholar with the University of Ottawa, where he completed the major content of his thesis. He has co-authored over 200 papers in established journals and conferences, and contributed to 13 book chapters. He is the founding director of Next Generation Communications and Computing Networks (NEXTCON)–Smart Connected Vehicles Innovation (SCVI) Labs. He has served as the Technical Program Co-Chair of more than ten international conferences/symposia/workshops. He is an Editor of the IEEE Communications Surveys & Tutorials and IEEE Transactions on Green Communications and Networking, an Associate Editor of IEEE Internet of Things Journal, IEEE Networking Letters, Vehicular Communications (Elsevier) and Sensors Journal. He also served as the Chair of the IEEE ComSoc Communication Systems Integration and Modeling Technical Committee. He is a Distinguished Speaker of the ACM, senior member of the IEEE and senior member of ACM.

Massimo Tornatore received the Ph.D. degree from Politecnico di Milano in 2006, where he is currently an Associate Professor with the Department of Electronics, Information, and Bioengineering. He also holds an appointment as an Adjunct Professor with the University of California at Davis, Davis, USA, and as a Visiting Professor with the University of Waterloo, Canada. His research interests include performance evaluation, optimization and design of communication networks (with an emphasis on the application of optical networking technologies) cloud computing, and machine learning application for network management. He has co-authored more than 400 peer-reviewed conference and journal papers (with 18 Best Paper Awards), 2 books and 1 patent, in the above areas. He is an Active Member of the Technical Program Committee of various networking conferences, such as Infocom, OPC, ICC, and Globecom. He is a member of the Editorial Board of the IEEE Communication Surveys & Tutorials, IEEE Communication Letters, Photonic Network Communications (Springer), and Optical Switching and Networking (Elsevier).