Data Center Network Placement and Data Backup Against Region Failures

by

Lisheng Ma

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Graduate School of Systems Information Science)
in Future University Hakodate
September 2017

To my family

ABSTRACT

Data Center Network Placement and Data Backup Against Region Failures

by

Lisheng Ma

Rapid growth of cloud computing has enabled a wide scope of new applications such as e-commerce and social networking. As the underlying supporting infrastructure, data center networks (DCNs) deployed in geographically distributed (geo-distributed) locations are becoming increasingly important. However, geo-distributed DCNs are vulnerable to large-scale region failures due to disasters. This makes DCN protection against region failures a critical task. Proactive protection is an important way to fight against DCN failures by network planning before disasters occur. To this end, this thesis investigates DCN placement and data backup against region failures via proactive protection mechanisms.

We first study optimal DCN and content placement with the objective of minimizing DCN failure probability. In this part, we combine the probabilistic region failure model and the grid partition scheme to capture the key features of the general non-uniform distribution of a potential region failure (in terms of its occurring probability and intensity) and to conduct network vulnerability assessment. Based on the vulnerability information, we further develop an integer linear program (ILP)-based theoretical framework to achieve optimal DCN and content placement with the mini-

mum DCN failure probability. A heuristic is also proposed to make our solution more scalable for large-scale networks.

We then optimize data backup for a particular DCN node threatened by an upcoming disaster by properly exploring the ε early warning time of the disaster, where ε denotes the time interval between the earliest moment that the DCN node is aware of the disaster and the latest moment that the disaster indeed hits the DCN. In this part, we investigate urgent data backup within the ε early warning time of the disaster for both homogeneous and heterogeneous data backup scenarios (the former concerns with the scenario that different types of data are backed up to the same set of backup DCN nodes while the latter considers the scenario that different types of data may be backed up to the different sets of backup DCN nodes).

In the homogeneous data backup scenario, we divide our design into two subproblems: Backup Capacity Evaluation (BCE) and Backup Cost Minimization (BCM). BCE helps DCN operators to find the maximum backup capacity, and thus fully utilize the early warning time to back up as much data as possible. Since the maximum backup capacity may not be sufficient for backing up all data, priority can be given to those more important data. On the other hand, BCM minimizes backup cost by properly selecting a set of safe backup DCN nodes and routes for those more important data. We propose both ILPs and heuristic for the two sub-problems.

In the heterogeneous data backup scenario, we propose two backup schemes: maximum data backup scheme (MDBS) and fairness data backup scheme (FDBS). The former maximizes the total amount of data that can be backed up, and the latter maximizes the same proportion of data backup for each type of data in a fair manner. For each scheme, an ILP and a heuristic are proposed to properly select a set of safe backup DCN nodes and corresponding backup routes.

Our proposed solutions for DCN and content placement can effectively protect DCNs and contents against a potential region failure under the global non-uniform

distribution. By taking the early warning time into account, our proposed backup schemes can generate efficient solutions for urgent data backup against ε-time early warning disaster. It is expected that the study in this thesis can provide a fundamental guideline to the design of disaster survivable DCNs.

ACKNOWLEDGEMENTS

I would like to extend my sincere gratitude to all those who helped me during my Ph.D. study in Future University Hakodate. Without your great help this thesis would not have been reached its present form.

First of all, I would like to express my deepest gratitude to my supervisor Professor Xiaohong Jiang, for his careful guidance and constant support in my academic research. He has spent great efforts to discuss the research topics with me and teach me a lot of skills in academic research and truth in life. Working with him is proved to be a very enjoyable and rewarding experience. I would also like to express my heartfelt gratitude to Professor Jiang's wife, Mrs Li, for her countless care.

I would like to thank my thesis committee members, Professor Yuichi Fujino, Professor Hiroshi Inamura and Professor Masaaki Wada for their valuable comments that help me a lot in improving both the quality and the clarity of this thesis.

My thanks would go to Professor Bin Wu of Tianjin University, China, Professor Achille Pattavina of Politecnico di Milano, Italy, Professor Norio Shiratori of Waseda University, Japan and Professor Tarik Taleb of Aalto University, Finland for their constructive suggestions in my Ph.D. studies. I also want to thank other members in our laboratory, all other teachers and university staffs for their help in these three years. It is a wonderful experience for me to study in Future University Hakodate.

I would like to express my sincere thanks to my beloved family for their unconditional love and great support in my life. I also owe a special debt of gratitude to my wife for understanding, encouragement and support in these three years.

TABLE OF CONTENTS

CHAPTER

LIST OF FIGURES

Figure

LIST OF TABLES

Table

CHAPTER I

Introduction

In this chapter, we first introduce the background of data center networks and disaster threats. Then we describe the motivations and contributions of this thesis. Finally, we give the outline of this thesis.

1.1 Data Center Networks

A data center network (DCN) is a warehouse-scale and massively parallel computing and storage resource. It consists of hundreds or even thousands of servers organized in racks, which are connected with a high-speed communication network [1–3]. In recent years, many large enterprises (e.g., Google, Amazon and Microsoft) have built their own DCNs in geo-distributed locations around the world to provide cloud services [4–6]. For example, according to [7], Google has more than 30 data centers around the world which include more than 450,000 servers and can process more than 20 petabytes of data per day.

Nowadays, most online services are geo-distributed to serve millions of users around the world such as online video, social networking, web search, etc., and then geo-distributed DCNs make it easy for any service to become geo-distributed [8, 9]. Based on geo-distributed DCNs, services or contents can be replicated among multiple DCNs located at different network regions and services can be provided by the anycast service mode (i.e., a service request can be served by any DCN that contains such service.) [10]. Such geo-distributed DCNs bring the following benefits: 1) a service request can be served by a nearby DCN that provides such service such that the service cost and latency can be reduced; 2) they can improve service survivability under failures, as services can still be supported by other DCNs containing the replicas of these services upon the failures of services at a particular DCN; 3) they can reduce the operating cost by exploiting the regional differences in prices of energy and real estate.

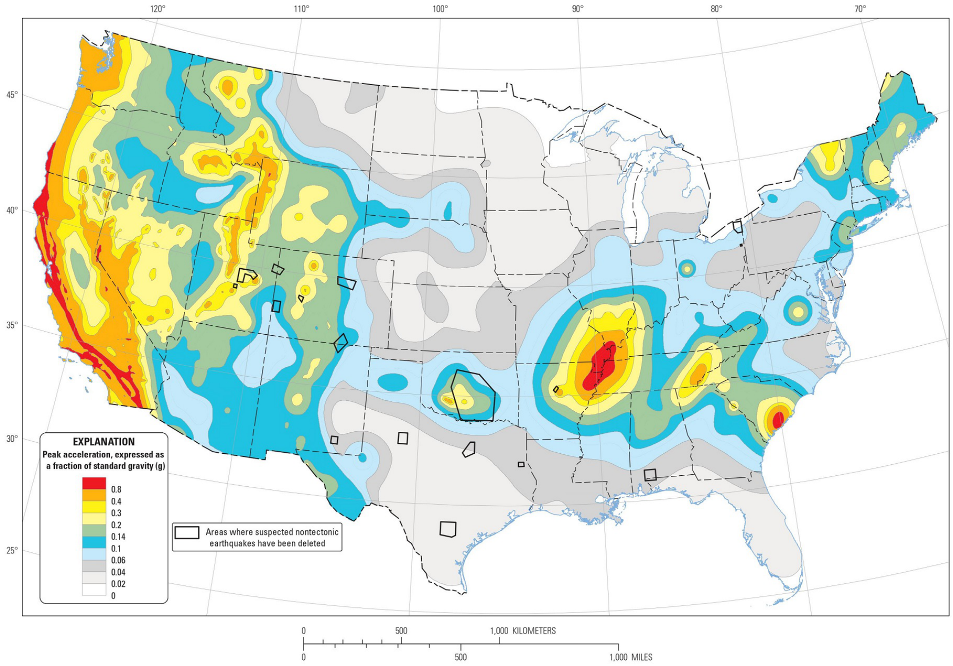
Due to these attractive advantages of geo-distributed DCNs, they are becoming important infrastructures to meet the growing demands of the emerging applications such as e-commerce, social networking, cloud computing, etc. As the trends like our increasing reliance on online services and many applications in mobile device changing into cloud services develop, it is believe that geo-distributed DCNs will play a more important role in the future communication networks.

1.2 Disaster Threats

With the increase of frequencies of disasters, geo-distributed DCNs are facing more and more potential large-scale disaster threats, both natural and human-made. Some recent major network disruptions due to disasters include 2012 Sandy Hurricane, 2011 Japan Tsunami, 2008 China Wenchuan earthquake, etc. [11–19]. Such disasters usually affect a specific geographical region, causing failures of a set of network components and degradations or even breakdowns of vital network services. For example, China Wenchuan earthquake in 2008 affected over 60 enterprise DCNs and more than 3000 telecom offices, as well as around 30,000 kilometers optic cables [13, 18], and Japan Tsunami and earthquake in 2011 affected tens of DCNs and more than 2000 telecom buildings [15, 19].

It is notable that different disasters with different features (e.g., intensity, predictability and location) lead to different impacts on network. Thus, network operators should consider different measures for different types of disasters to protect network. For the natural disasters such as earthquakes, hurricanes, floods and tsunamis, based on climatic and environmental conditions the potential intensity and location of those disasters can be estimated by using predictable technologies of disasters [20, 21] before disasters occur. Then, network operators can take the potential disasters into account in the network planning stage (e.g., deploying a new DCN). On the other hand, the early warning systems for disasters are widely applied in the world which can help us to obtain certain early warning information (e.g. affected region and time) of an upcoming disaster. For example, REIS (real-time earthquake information system) [22] is an earthquake early warning system deployed in Japan. It can estimate location and magnitude of an earthquake within 5 seconds after the P-waves arrive. Besides, national hurricane center in America [21] can provide early hurricane warnings on a time basis from hours to days. For different types of disasters, we can obtain different early warning times (from a few seconds to a few days). Based on the early warning information, network operators can carry out the urgent protection schemes for the network facilities that will be affected by an upcoming disaster.

In addition to natural disasters, human-made disaster threats such as weapons of mass destruction (WMD) attacks, electromagnetic pulse (EMP) attacks are rising [23, 24]. In general, human-made attacks choose large cities and important infrastructures as targets such as government, DCNs. Thus, network operators also need to consider the possible human-made disasters in the network deployed regions when they design the network protection schemes.


120°
110°
100°
90°
80°
70°
45°
40°
35°
30°
25°
EXPLANATION
Peak acceleration, expressed as a fraction of standard gravity (g)
0.8
0.4
0.3
0.2
0.14
0.1
0.06
0.04
0.02
0
Areas where suspected nontectonic earthquakes have been deleted
0
500
1,000 KILOMETERS
0
500
1,000 MILES

Two-percent probability of exceedance in 50 years map of peak ground acceleration

Figure 1.1: U.S. national seismic hazard map

1.3 Motivations and Contributions

As the above discussions, geo-distributed DCNs are vulnerable to large-scale disaster threats. Thus, it is crucial to study the DCN protection measures against region failures due to disasters, and then disaster survivable DCNs can be achieved [25–30]. To this end, this thesis focuses on the DCN and content placement with the consideration of a potential large-scale region failure due to disaster and data backup in DCNs against an upcoming disaster. Given a network, the DCN and content placement in the network with the consideration of a potential region failure usually concerns with the following two aspects: 1) to assess the network vulnerability due to a region failure; 2) based on the network vulnerability information, to properly place the DCNs and contents in the network such that the DCN failure probability due to region failure is minimized. For network vulnerability assessment, the previous works [31–35] all assumed that both occurring probability and intensity of region failure(s) follow the uniform distribution in the network area (Please see Section 2.1 for related work). As illustrated in Fig. 1.1, from U.S. national seismic hazard map [20], we can observe that in the real world, however, a disaster may happen in different areas with different probabilities and different intensities (i.e., non-uniform distribution of a disaster). Thus, it is desirable to capture the key features of the general non-uniform distribution of a potential region failure due to disaster in terms of its occurring probability and intensity, and then apply them to conduct the network vulnerability assessment.

Note that since DCN and content placement with the consideration of a potential region failure is implemented based on network vulnerability information, the previous works [36–40] on DCN and content placement also failed to take into account the global non-uniform distribution of potential region failures in terms of their occurring probabilities and intensities (Please see Section 2.2 for related work). On the other hand, in a large-scale network there are multiple paths between an arbitrary pair of nodes, which indicates that the probability that these paths simultaneously fail due to disaster is very small. In contrast, if a DCN hosting node fails after disaster, the contents provided by this node will be unavailable and the adverse impact of such failure on the DCN is even greater than the path failure. Thus, the tradeoff between failure probabilities of DCN hosting nodes and failure probabilities of requesting paths should be considered. Also, since content or service providers in DCNs wish to satisfy user demands with low latency, we need to consider the traffic transmission delay issue as well in the DCN design.

Data backup is an important proactive approach against disasters in DCNs by storing multiple redundancies across geo-distributed DCNs. The existing studies mainly focused on periodical data backup [41, 42] (Please see Section 2.3 for related work). Such periodical backup schemes may not result in high data protection efficiency under the disaster scenario, because a sudden disaster generally occurs in an unpredictable manner, and thus newly generated data may not be well protected in time due to the fixed data backup period. Based on the early warning information from the early warning systems, recently early warning time backup against disasters was proposed in [43] and [44] to maximize data owners' utility and the number of contents that can be evacuated, respectively. However, this problem have not been fully explored yet. For example, backup cost as a major concern for DCN operators to select a protection strategy and the heterogeneous data backup scenario (i.e., different types of data hosted at a DCN node may be backed up to the different sets of backup DCN nodes) are not considered.

To address the above limitations on the DCN placement and data backup against region failures due to disasters, this thesis studies the DCN and content placement with the consideration of global non-uniform distribution of a potential region failure and urgent data backup by fully utilizing the early warning time of an upcoming disaster. The main contributions of this thesis are summarized as follows.

1. Region failure-aware DCN and content placement.

We study the optimal DCN and content placement in this part to minimize the DCN failure probability under a region failure. We first propose a general grid partition-based scheme to evaluate the vulnerability of a given network due to the global non-uniform distribution of a region failure, in which the probabilistic region failure model is applied to determine the failure probability of a node/link. Then we can create a "vulnerability map" for DCN and content placement in the network. Based on the grid partition-based scheme and the corresponding vulnerability map, we further develop an integer linear program (ILP)-based theoretical framework to achieve optimal DCN and content placement, which leads to minimum DCN failure probability against a region failure. To make the problem more scalable for large-scale problems, a heuristic is also proposed to achieve the time-efficient solution. Finally, we present extensive numerical results to demonstrate the validity of the proposed network vulnerability assessment scheme and the proposed ILP and heuristic for DCN and content placement.

2. Homogeneous data backup based on early warning of region failure.

In this part, we investigate the urgent data backup for a particular DCN node threatened by a region failure due to an upcoming disaster with the early warning time ε, where we consider the homogeneous data backup (i.e., different types of data hosted at the DCN node are backed up to the same set of backup DCN nodes). We first formulate an ILP to find the maximum amount of data that can possibly be protected by fully utilizing the given early warning time ε. This helps to determine which data should be protected according to data importance. Then, we formulate another ILP to achieve minimum cost backup by properly selecting a set of safe backup DCN nodes and corresponding backup routes for those selected important data. To get real-time solutions for engineering practice, we also propose a heuristic to achieve cost-efficient backup for ε-time early warning disaster. Finally, extensive numerical results show that our solutions can be self-adaptive to different early warning times.

3. Heterogeneous data backup based on early warning of region failure.

In this part, we also focus on the optimal data backup for a particular DCN node threatened by a region failure due to an upcoming disaster with the early warning time ε, where the heterogeneous data backup (i.e., different types of data hosted at the DCN node may be backed up to the different sets of backup DCN nodes) is taken into account. To this end, two backup schemes are developed to carry out urgent backup within the given early warning time ε, which are maximum data backup scheme (MDBS) and fairness data backup scheme (FDBS). The former is to maximize the total amount of data that can be backed up, and the latter is to maximize the same proportion of data backup for each type of data in a fair manner. For those backup schemes, we first develop the corresponding ILP models by properly selecting a set of safe backup DCN nodes and routes to obtain the optimal backup solutions. To meet the real-time requirement of engineering practice, we then propose the corresponding heuristics. Finally, extensive numerical results show that the solutions from both schemes are adaptive to different early warning times

1.4 Thesis Outline

The rest of this thesis is organized as follows. Chapter II discusses the related work of this thesis. We investigate the region failure-aware data center network and content placement in Chapter III. Chapter IV presents the work on homogeneous data backup based on early warning of region failure and Chapter V introduces the work regarding heterogeneous data backup based on early warning of region failure. Finally, we conclude this thesis, and discuss the topics for future research in Chapter VI.

CHAPTER II

Related Work

In this chapter, we present the previous works related to our study in this thesis, including network vulnerability assessment, data center network and content placement, as well as network protection.

2.1 Network Vulnerability Assessment

To evaluate network vulnerability under disasters, different models can be adopted to capture the key features of region failures due to disasters, which include deterministic model and probabilistic model [45]. Under deterministic model, any network component (e.g., node, link, etc.) fails with the probability 1 if it falls within the failure region due to a disaster, whereas that falling within the failure region fails with a certain probability between 0 and 1 based on probabilistic model, and such a failure probability depends on the intensity of failure, the distance to failure center and also the dimension of the component (such as the length of a link).

Based on the aforementioned region failure models, some works have been done on the assessment of network vulnerability and identification of vulnerable network zones due to region failure [31–35]. By using the deterministic circular/line cut region failure models, the network vulnerability assessments were conducted in [31, 32]. Since under a real-world disaster the network components rarely are completely destroyed, the real-world disasters have probabilistic rather than deterministic impacts on network components, and then probabilistic model is more suitable for network vulnerability assessment under disasters. In [33] and [34], a probabilistic failure model and grid partition based framework were developed to efficiently evaluate the network vulnerability. Recently, network vulnerability assessment with the consideration of multiple simultaneous probabilistic failures was investigated in [35].

In the above works, both occurring probability and intensity of region failure(s) follow the uniform distribution in the network area which cannot match the real-world disasters that may occur in different regions with different probabilities and different intensities. Thus, this thesis studies the network vulnerability assessment with the consideration of the global non-uniform distribution of a potential region failure due to disaster in terms of its occurring probability and intensity.

2.2 Data Center Network and Content Placement

Regarding the data center network (DCN) and content placement, the work in [46] studied DCN and content placement with the objective of minimizing the network's power consumption. To solve the scalability issue, a fully scalable DCN architecture with distributed placement of component sets in a given optical network was proposed in [47], which can remove the environmental constraints and also reduce the system cost. Recently, content placement was considered in [48] to identify the optimal placement of videos in a large-scale VoD system such that the total network bandwidth consumption is minimized.

With the consideration of potential network failure(s), Xiao *et al.* in [36] studied the optimal DCN placement problem with service routing and protection to minimize the network cost, while ensuring fast protection of all services against any single link failure or service failure at a particular DCN. By assuming multiple region failures in fixed locations, the works in [37, 38] concerned with the joint design of content placement, routing, and protection of paths and contents to achieve more efficient protection of optical DCNs than dedicated single-link failure protection, while the works in [39, 40] investigated the DCN and content placement to minimize both the contents unavailability due to DCN hosting nodes damage and requests unreachability due to paths damage from disasters. Besides, extensive efforts have been focused on node placement problems considering minimizing the traffic weighted mean internodal distance of a network, the number of deployed nodes, cost, etc. [49–51].

The previous works on the DCN and content placement with the consideration of potential network failure(s) failed to take into account the non-uniformly distributed region failure, and these works also did not consider the inherent tradeoff among failure probabilities of DCN hosting nodes, failure probabilities of requesting paths and traffic transmission delay. This thesis investigates the DCN and content placement with the consideration of global non-uniform distribution of a potential region failure, where the tradeoff among failure probabilities of DCN hosting nodes, failure probabilities of requesting paths and traffic transmission delay is considered.

2.3 Network Protection

Network protection against region failures due to disasters can be achieved by either proactive approaches or post-disaster restoration schemes [52]. The former designs scheme to prevent network failures by network planning before disasters occur. The latter utilizes resources available at the disaster time to recover network. Due to the uncertainty of disasters, proactive approaches require a relatively large amount of resources to achieve a desired level of protection. In contrast, post-disaster restoration is cost-saving, but the effect is generally poor due to the best-effort nature. For proactive approaches, the work in [38] studied the protection scheme against a single disaster failure by providing the backup path and data center for a request affected by the disaster. A disaster-risk-aware provisioning was proposed in [53] in which valuable connections are routed on no-(or low-) risk regions due to disasters such that the risk and penalty can be reduced under such disasters, and the link-disjoint primary and backup paths are also provided to avoid the simultaneous failures of those paths under disasters. The study in [54] focused on the disaster-aware service provisioning scheme which multiplexes service over multiple paths destined to multiple serves (or data centers) with manycasting against failures of links and nodes caused by disasters.

In terms of proactive approach, data backup is an important proactive protection method. Based on the mutual backup model in [55], some periodical data backup schemes were proposed in [41] and [42] to jointly optimize backup site selection and data transmission paths. Recently, early warning time backup against disasters was proposed in [43] and [44] to carry out urgent backup within the early warning time for those data that may not be well protected by regular backup in time due to the fixed data backup period. The work in [43] evacuated as much contents as possible from the DCN node threatened by disaster to a single backup DCN node within the early warning time, while the study in [44] carried out time-constrained urgent backup to maximize data owners' utility. In addition, some works [56] and [57] focused on the real-time data replications in DCNs whereas data generated in a certain past period of time is not considered.

Regarding post-disaster restoration schemes, three post-disaster reprovisioning schemes were proposed in [58] to maintain network connectivity and maximize the traffic flow in the post-disaster network. The work in [59] considered the issue of restoration in optical cloud networks for fiber link failure and then a restoration-based survivability strategy was developed by combining the benefits of both cloud service relocation and service differentiation concepts to restore cloud services. A post-disaster re-provisioning scheme was proposed for telecom mesh networks in [60] which takes fairness-aware degradation and multipath deployment into account. Under a large-scale disaster, multiple network components will be affected, and then the failed components may be repaired through multiple restoration stages for some reasons e.g., limited repair resources. Thus, the progressive disaster recovery is an attractive topic in recent years which was investigated in [61–63]. Some summaries of network protections against disasters were presented in [64–68].

In this thesis, the study of DCN protection falls into the same category as [43] and [44]. We propose urgent backup schemes for homogeneous and heterogeneous data backup, respectively, and our solutions can be self-adaptive to different early warning times.

CHAPTER III

Region Failure-Aware Data Center Network and Content Placement

In this chapter, we focus on the region failure-aware data center network (DCN) and content placement in which the non-uniform distribution of a potential region failure due to disaster is considered. Given a network for DCN placement, a general probabilistic region failure model is adopted to capture the key features of a region failure and to determine the failure probability of a node/link in the network under the region failure. We then propose a general grid partition-based scheme to flexibly define the global non-uniform distribution of a potential region failure in terms of its occurring probability and intensity. Such grid partition scheme also helps us to evaluate the vulnerability of a given network under a region failure and thus to create a "vulnerability map" for DCN and content placement in the network. With the help of the vulnerability map and by taking into account the tradeoff among failure probabilities of DCN hosting nodes, failure probabilities of requesting paths and traffic transmission delay, we further develop an integer linear program (ILP)-based theoretical framework to identify the optimal DCN and content placement, which leads to the minimum DCN failure probability against a region failure. To make the overall placement problem more scalable for large-scale networks, a heuristic is also proposed by dividing the problem into two sub-problems (i.e., DCN placement and content placement). Finally, extensive numerical experiments are carried out based on the real gridded data of U.S. national seismic hazard map [69] to demonstrate our proposed network vulnerability assessment scheme and to validate the efficiency of the proposed ILP and heuristic for DCN and content placement under non-uniform spatial and intensity distribution of a potential disaster.

3.1 Network Vulnerability Assessment

We consider a network with deployment area Z and denote it as a graph $G = (V, E)$, where V is a set of nodes and E is a set of network links.

3.1.1 Probabilistic Region Failure Model

A real-world disaster is usually confined in a specific geographical region. A network component (like a link or node) in this disaster region will fail with certain probability, and such a failure probability depends on the intensity of failure, the distance to failure center and also the dimension of the component (such as the length of a link). To capture these key features of a region failure, we adopt the general probabilistic region failure (PRF) model proposed in [34].

- ***PRF Model Definition*:**

(1) As illustrated in Fig. 3.1, the PRF is defined by a set of consecutive concentric annuluses with radius $r_i, i = 1, ..., m$.

(2) The ith annulus is associated with failure probability p_i, and such probability is monotonously decreasing with annulus, i.e., $p_i \geq p_{i+1}$, $1 \leq i \leq m-1$. Here, the region failure is only confined within the circle area of radius r_m, beyond which the failure probability is regarded as 0.

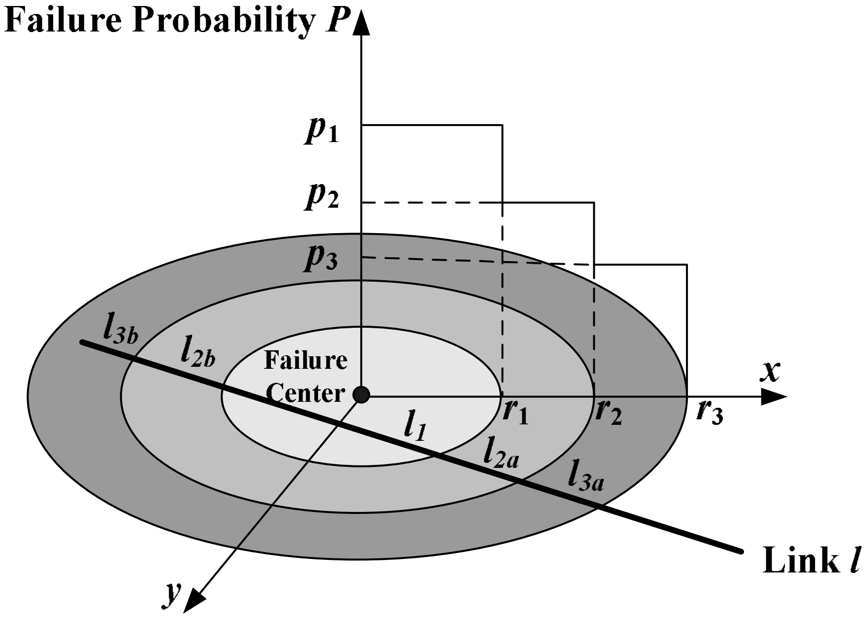


Figure 3.1: Probabilistic region failure model, m=3

It is notable that under a probabilistic region failure, multiple network components (e.g. nodes and links) may simultaneously fail, but with a certain probability for each. In this thesis we evaluate failure probabilities of node and link separately without any dependency between the two. Since failure probability evaluations of nodes and links are different from each other as follows, the proposed approaches can properly handle various scenarios.

Based on the PRF model, the failure probability P_v for a node v in the ith annulus can be formulated as

$$P_v = p_i. \tag{3.1}$$

In general, a link spans multiple annuluses of a region failure, and each annulus contains a segment of the link. Then, failure probability of the link is determined by that of all those segments. Therefore, the failure probability P_l for a link l can be formulated as

$$P_l = 1 - \prod_{i=1}^{m}(1 - P_{l_i}), \tag{3.2}$$

where m is the number of annuluses in the PRF model and P_{l_i} is the failure probability of segment l_i on link l that falls into the ith annulus.

Consider a segment l_i on link l that falls into the ith annulus. We first divide such a segment into multiple shorter segments, and each of them is approximated as a node to evaluate the failure probability of l_i. Then, the failure probability P_{l_i} for l_i can be formulated as

$$P_{l_i} = 1 - (1 - p_i)^{\frac{|l_i|}{\xi}}, \tag{3.3}$$

where ξ is a pre-defined factor representing the length of the shorter segment and $|l_i|$ represents the length of segment l_i. Note that in a practical fiber-optical network, each fiber link has a set of amplifiers. Generally, a link failure is mainly caused by failures of those amplifiers. Similar to [35], we can equivalently treat a segment on a particular link as a sequence of amplifiers, with each approximated as a node to evaluate its failure probability. This explains equation (3.3).

For the example in Fig. 3.1, the failure probabilities of segments on link l are evaluated as

$$\begin{aligned} P_{l_1} &= 1 - (1 - p_1)^{\frac{|l_1|}{\xi}}, \\ P_{l_2} &= 1 - (1 - p_2)^{\frac{|l_2|}{\xi}}, \\ P_{l_3} &= 1 - (1 - p_3)^{\frac{|l_3|}{\xi}}, \end{aligned} \tag{3.4}$$

where

$$|l_2| = |l_{2a}| + |l_{2b}|, |l_3| = |l_{3a}| + |l_{3b}|. \tag{3.5}$$

Based on link failure probability, failure probability P_r for a path r can be formulated as

$$P_r = 1 - \prod_{l \in r}(1 - P_l), \tag{3.6}$$

where P_l is the failure probability of a link l on path r.

3.1.2 Vulnerability Metrics

To evaluate the vulnerability of a network, we consider the following two vulnerability metrics:

- **NFP (node failure probability):** The probability that a node fails due to a PRF.
- **LFP (link failure probability):** The probability that a link fails due to a PRF.

For a given network, one straight-forward approach to assessing the vulnerability of a metric $\triangle$ is to first partition the overall network area into some disjoint region failure location (RFL) zones

- ***RFL Zone Definition*:** A RFL zone for a specified metric $\triangle$ (e.g. NFP or LFP) is a network subarea that any PRF with center in it will always induce the same value of $\triangle$ to the network.

For a specified metric $\triangle$, suppose that we have already divided the network deployment area Z into a set of disjoint RFL zones Z_n, where a PRF in Z_n induces the value $\triangle_{Zn}$ of $\triangle$ to the network. Then the overall metric $\triangle$ can be calculated as

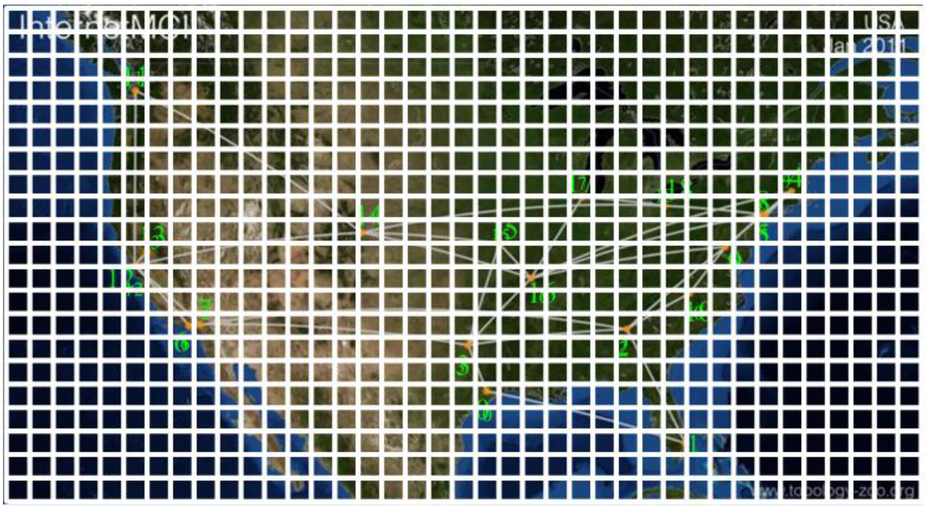


Figure 3.2: A grid partition for U.S. InternetMCI network

$$\triangle = \sum_{Z_n} P_{Z_n} \cdot \triangle_{Z_n}. \tag{3.7}$$

Here, P_{Z_n} denotes the probability that a PRF falls within the RFL zone Z_n.

It is notable that to directly apply (3.7) for calculating a metric $\triangle$, we first need to find out all RFL zones of the metric, which involves the complicated geometric computation and quickly becomes computationally intractable for a large-scale network [33, 34]. In the following section, we propose a general grid partition-based scheme, which helps us to flexibly define the non-uniform distribution of PRF and to efficiently evaluate the vulnerability of a network.

3.1.3 Grid Partition-Based Vulnerability Evaluation

As illustrated in Fig. 3.2 for U.S. InternetMCI network [70], we apply a grid partition scheme to evenly divide the network area Z into M small square cells. Based on this grid partition scheme, if we regard each cell as a "RFL" zone and take the center point of the cell as the failure center to calculate the metric $\triangle$, then we

Algorithm 1 NFP evaluation:

```
Input:
    Network topology information, a set of nodes $V$ and failure model parameters.
Output:
    $NFP$: $\triangle_{NFP_v}$ evaluation for node $v \in V$.
for each node $v$ in $V$ do
    $\triangle_{NFP_v} = 0$;
    for $n \in [1, 2, ..., M]$ do
        calculate $NFP$ $\triangle^{v}_{Z_n}$ for $v$ by using $(x_{Z_n}, y_{Z_n})$ as the center point of concentric
        circles in PRF model with parameters $Z^{para}_{n}$;
        $\triangle_{NFP_v} = \triangle_{NFP_v} + P_{Z_n} \cdot \triangle^{v}_{Z_n}$;
    end for
end for
return $\triangle_{NFP_v}, v \in V$.
```

Algorithm 2 LFP evaluation:

```
Input:
    Network topology information, a set of links $E$ and failure model parameters.
Output:
    $LFP$: $\triangle_{LFP_l}$ evaluation for link $l \in E$.
for each link $l$ in $E$ do
    $\triangle_{LFP_l} = 0$;
    for $n \in [1, 2, ..., M]$ do
        calculate $LFP$ $\triangle^{l}_{Z_n}$ for $l$ by using $(x_{Z_n}, y_{Z_n})$ as the center point of concentric
        circles in PRF model with parameters $Z^{para}_{n}$;
        $\triangle_{LFP_l} = \triangle_{LFP_l} + P_{Z_n} \cdot \triangle^{l}_{Z_n}$;
    end for
end for
return $\triangle_{LFP_l}$, $l \in E$.
```

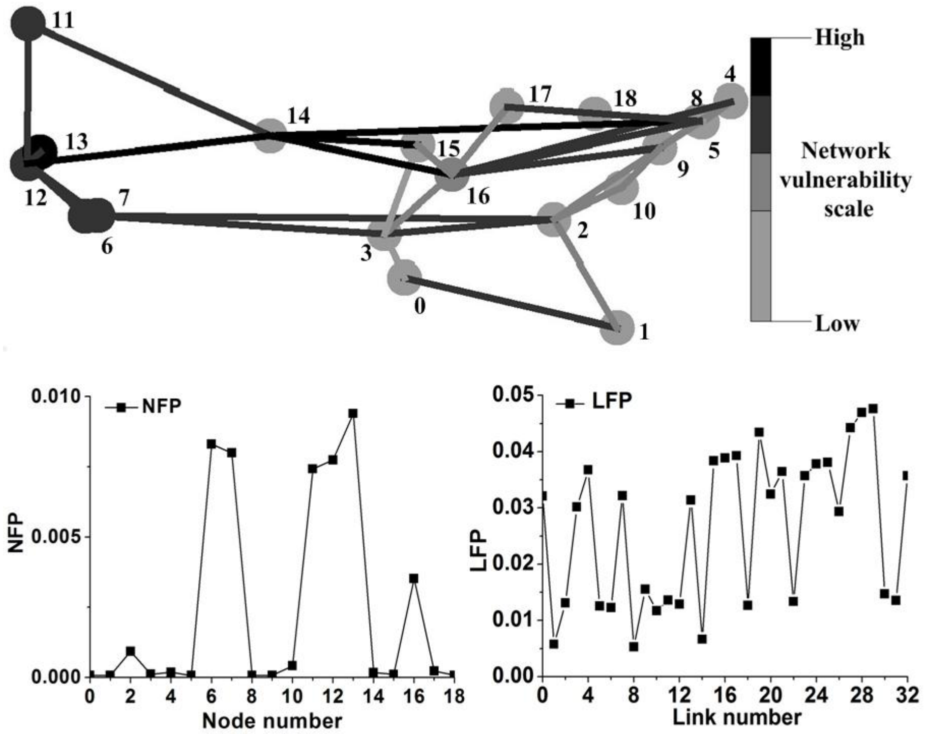


Figure 3.3: Vulnerability map

can get an evaluation of metric $\triangle$ based on (3.7). Since the intensity of a disaster may be different in different regions, a PRF with center falling within different cells may have different parameters of r_i and p_i.

If we use (x_{Z_n}, y_{Z_n}) to denote the center point of cell Z_n, with the help of the grid partition scheme the evaluations of NFP and LFP are summarized as Algorithms 1 and 2, respectively. Here, the number of square cells M, PRF model parameters Z_n^{para} and the probability P_{Z_n} that a PRF falls within the zone Z_n can be determined according to the information of real disaster data, such as the gridded data of U.S. national seismic hazard map [69].

It is notable that the grid partition scheme can also help us to create a "vulnerability map" of a given network, in which the NFP for each node and LFP for each link in the network are illustrated.

For example, for the network shown in Fig. 3.2, its "vulnerability map" is shown in Fig. 3.3 (See Table 3.1 for link information and subsection 3.4.1 for related parameter

Table 3.1: Links in the U.S. InternetMCI network

Link No.	Link	Link No.	Link	Link No.	Link
0	(0,1)	11	(4,8)	22	(9,10)
1	(0,3)	12	(4,9)	23	(9,16)
2	(1,2)	13	(4,16)	24	(11,12)
3	(2,3)	14	(5,8)	25	(11,14)
4	(2,7)	15	(6,7)	26	(12,13)
5	(2,9)	16	(6,12)	27	(12,14)
6	(2,10)	17	(7,12)	28	(14,15)
7	(3,7)	18	(8,9)	29	(14,16)
8	(3,15)	19	(8,14)	30	(15,16)
9	(3,16)	20	(8,16)	31	(16,17)
10	(4,5)	21	(8,18)	32	(17,18)

settings). Such “vulnerability map” will be helpful for identifying the optimal DCN and content placement in the network to lead to the minimum DCN failure probability.

3.2 ILP for Data Center Network and Content Placement

With the help of the “vulnerability map” of a given network, we consider here the optimal DCN and content placement in the network to minimize the DCN failure probability due to a region failure. The inherent tradeoff among failure probabilities of DCN hosting nodes, failure probabilities of requesting paths and traffic transmission delay is also considered in the optimal placement problem.

3.2.1 Problem Description

In this placement problem, we consider to place multiple DCNs and different types of contents in a given network, and each DCN and each type of content are treated equally. Our objective is to determine the locations of DCNs and contents in a given network such that DCN failure probability under a region failure is minimized. Thus, we consider the simple scenario in which the size of each type of content and the constraints of bandwidth on each link, storage capacity of each DCN deployed node and the service ability of each DCN node are not taken into account. Regarding a request for a content, we only consider the traffic transmission delay to avoid the long communication latency between the requesting node and content hosting node, and other requirements are also not taken into account. We use the length of a path to approximate the transmission delay of the traffic along it, and formulate the optimal DCN and content placement problem as an ILP problem as follows.

3.2.2 Notation List

The detailed inputs and variables used in the ILP formulation are listed in Tables 3.2 and 3.3.

Table 3.2: Parameters for Inputs

Notation	Definition
V	The set of all nodes in network G(V, E).
E	The set of all links in network G(V, E).
V'	The set of DCN candidate hosting nodes, $V^{'} \subseteq V$.
C	The set of contents provided by DCNs.
δ	The scaling factor for adjusting the weight among total failure probability of DCN hosting nodes, total failure probability of requesting paths and total traffic transmission delay.
S	The set of requesting nodes, $S \subseteq V$.
R_{sv}	The set of paths between requesting node s and DCN hosting node v.
N_d	The number of DCNs to be placed.
N_c	The maximum number of replicas of content c.

N_{sv}	The number of paths between requesting node s and DCN hosting node v.
β	Predefined constant greater than the number of contents $\|C\|$.
PF_v	The failure probability of DCN candidate hosting node v ($\triangle_{NFP_v}$) obtained by "vulnerability map".
PF_{rsv}	The failure probability of path r between requesting node s and DCN hosting node v obtained by $P_r = 1 - \prod\limits_{l \in r}(1 - P_l)$.
PF_{sv}	The average failure probability of paths between requesting node s and DCN hosting node v.
L_{rsv}	The length of path r between requesting node s and DCN hosting node v.
L_{sv}	The average length of paths between requesting node s and DCN hosting node v.

Table 3.3: Variables

Notation	Definition
H_v	Binary variable. It takes 1 if a DCN is placed at node v and 0 otherwise.
H_v^c	Binary variable. It takes 1 if content c is hosted at DCN hosting node v and 0 otherwise.
H_v^{sc}	Binary variable. It takes 1 if requesting node s requests content c provided by DCN hosting node v and 0 otherwise.

3.2.3 ILP Formulation

$$Minimize \quad \Big\{\delta \sum_{v \in V'} H_v PF_v + \sum_{v \in V'} \sum_{s \in S} \sum_{c \in C} H_v^{sc}(PF_{sv} + L_{sv})\Big\}. \tag{3.8}$$

Subject to

$$PF_{sv} = \frac{\sum_{r \in R_{sv}} PF_{rsv}}{N_{sv}}, \forall s \in S, \forall v \in V'; \tag{3.9}$$

$$L_{sv} = \frac{\sum_{r \in R_{sv}} L_{rsv}}{N_{sv}}, \forall s \in S, \forall v \in V'; \tag{3.10}$$

$$H_v \geq \frac{1}{\beta} \sum_{c \in C} H_v^c, \forall v \in V'; \tag{3.11}$$

$$\sum_{v \in V'} H_v \leq N_d; \tag{3.12}$$

$$\sum_{v \in V'} H_v^c \geq 2, \forall c \in C; \tag{3.13}$$

$$\sum_{v \in V'} H_v^c \leq N_c, \forall c \in C; \tag{3.14}$$

$$H_v^{sc} \leq H_v^c, \forall v \in V', \forall s \in S, \forall c \in C; \tag{3.15}$$

$$\sum_{v \in V'} H_v^{sc} = 1, \forall s \in S, \forall c \in C. \tag{3.16}$$

Objective (3.8) (abbreviated as $failure\,risk$) minimizes the total failure probability of DCN hosting nodes and requesting paths, as well as the total traffic transmission delay. The scaling factor δ is used to control the weight among total failure probability of DCN hosting nodes, total failure probability of requesting paths and total traffic transmission delay. Equation (3.9) determines the average failure probability of paths between requesting node s and DCN hosting node v while Equation (3.10) calculates the average length of paths between requesting node s and DCN hosting node v. Constraint (3.11) implies that if any content c is provided by a node v, then a DCN must be placed at this node. Here, we use β larger than $|C|$ to ensure that constraint (3.11) can be properly established when $H_v = 1$ and $1 \leq \sum\limits_{c \in C} H_v^c \leq |C|$. Constraint (3.12) indicates a bound on the total number of DCNs placed in the network. Constraint (3.13) guarantees that any content c is replicated at least twice while constraint (3.14) limits the number of replicas of content c to its maximum possible number. Constraint (3.15) ensures that if requesting node s requests content c provided by DCN hosting node v, node v should contain content c. Constraint (3.16) guarantees that a request from node s for content c can be satisfied by only one DCN containing content c.

In our work, DCN placement is static, which is implemented at the network planning stage for only once. However, since the information on disaster and content properties (e.g. content request) is time-varying, content placement can be adjusted when the information on disaster and content properties are updated. In general, content placement can be optimized either periodically according to daily content requests variation, or within the early warning time of an upcoming disaster if the DCN failure risk is observed higher than the current risk evaluation. It is notable that the content requests (i.e., connection requests from requesting nodes for contents) only depend on the requesting nodes and the amount of contents, and are independent from the final locations of DCN and content placement. As requesting nodes and contents are given, the content requests can be modeled/obtained based on those given parameters by simple statistics.

3.3 Heuristic

To make the overall placement problem more scalable for large-scale networks, we propose here a heuristic to divide the problem into two sub-problems. We first solve the DCN placement problem, and then consider the content placement problem by taking the results of DCN placement as the input.

3.3.1 Algorithm Description

The proposed heuristic is summarized in Algorithms 3 and 4. Algorithm 3 gives the pseudo code of DCN placement, and then based on the results of Algorithm 3, the content placement scheme is shown in Algorithm 4. Here, the notations PF_{sv}, L_{sv}, PF_v, N_d, N_c, V', C, S and δ are defined in Section 3.2.2, and let $|B|$ denote the number of elements in an arbitrarily given set B.

DCN placement: In order to determine DCN hosting nodes, we need to evaluate the failure risk of each candidate DCN hosting node and then the selected DCN hosting nodes induce small failure risk for connection requests. Since the connection requests for each content are given which are independent from the final locations of DCNs and contents, we can use these connection requests (i.e., the information of content requests) to evaluate the failure risk of each candidate DCN hosting n-ode. For DCN placement, first, we find a content that has the maximum number of connection requests from requesting nodes. For each DCN candidate hosting node, we calculate the failure risk when the content found before is provided by this node, respectively. Then, we determine a node that induces the minimum failure risk as the first DCN hosting node from the set of DCN candidate hosting nodes. After that we use the iteration to determine all the DCN hosting nodes. In each iteration, a DCN hosting node from the set of DCN candidate hosting nodes is found which induces the minimum failure risk when all connection requests are provided by the determined DCN hosting nodes and this node. For a given network for DCN placement, our algorithm can find the DCN hosting nodes from the set of DCN candidate hosting nodes. Then, the small failure risk is achieved if all connection requests from a set of requesting nodes are provided by these nodes.

Content placement: After determining the DCN hosting nodes, we then assign the contents to DCNs which should satisfy constraints (3.13) and (3.14) in Section 3.2.3. For each content, we first assign it to these DCN hosting nodes, which induce the minimum value of total failure probability of requesting paths and total traffic transmission delay for this content provided by these nodes. To satisfy constraints (3.13) and (3.14), for each content, we check the number of DCN hosting nodes which host this content. If a content doesn't satisfy constraint (3.13), we successively assign this content to DCN hosting node that doesn't contain this content until satisfying constraint (3.13), which induces the minimum value of total failure probability of requesting paths and total traffic transmission delay for this content. If a content doesn't satisfy constraint (3.14), we successively reduce the DCN hosting node containing this content until satisfying constraint (3.14). Compared with the original DCN hosting nodes containing this content, we ensure that the remaining nodes bring about the smallest gap in the value of total failure probability of requesting paths and total traffic transmission delay for this content.

In Algorithm 3, the initialization is shown in line 1. The content $c \in C$ is found which has the maximum number of connection requests from requesting nodes in line

Algorithm 3 DCN Placement (DP):

```
Input:
    $G(V,E)$, $V' \subseteq V$, $S \subseteq V$, $PF_{sv}$ and $L_{sv}$ for $\forall s \in S, \forall v \in V'$, $PF_v$ for $\forall v \in V'$, $C$,
    $\delta$, $N_d$ and $(sc) \in R_c$: the set of connection requests for content $c \in C$, $s \in S$.
Output:
    The set of DCN hosting nodes: $L$.
 1: $L = \varnothing$; $Risk_{min} = \infty$;
 2: $c$=$\arg_{c \in C}$ $\max\{|R_c|\}$;
 3: for each $v \in V'$ do
 4:     $Risk_v = \delta PF_v + \sum_{(sc) \in R_c, \forall s \in S}(PF_{sv} + L_{sv})$;
 5:     if $(Risk_v < Risk_{min})$ then
 6:         $Risk_{min} = Risk_v$; $u = v$;
 7:     end if
 8: end for
 9: $L = L \bigcup \{u\}$;
10: while $(|L| < N_d)$ do
11:     $Risk_{min} = \infty$;
12:     for each $v \in (V' - L)$ do
13:         $Risk_v = \sum_{(sc) \in R_c, \forall s \in S, \forall c \in C} min_{\forall u \in (L \bigcup \{v\})}(PF_{su}$
14:                     $+L_{su})$;
15:         $Risk_v = Risk_v + \delta \sum_{\forall w \in (L \bigcup \{v\})} PF_w$;
16:         if $(Risk_v < Risk_{min})$ then
17:             $Risk_{min} = Risk_v$; $v' = v$;
18:         end if
19:     end for
20:     $L = L \bigcup \{v'\}$;
21: end while
22: return $L$;
```

2. From lines 3-8, for each DCN candidate hosting node v, we calculate the failure risk for content c provided by this node, respectively, and find the node u that induces the minimum failure risk. The failure risk is obtained in line 4. The node u is determined as the first DCN hosting node in line 9. All the DCN hosting nodes are determined through the iteration in lines 10-21. In each iteration, we select a DCN hosting node v' from the DCN candidate hosting node set $V' - L$, and we can obtain the minimum failure risk when all connection requests are provided by these DCN hosting nodes in $L \bigcup \{v'\}$. Here, the failure risk is obtained in lines 13-15.

In Algorithm 4, we can implement the content placement to satisfy constraints

Algorithm 4 Content Placement (CP):

```
Input:
    $G(V,E)$, $V' \in V$, $S \in V$, $PF_{sv}$ and $L_{sv}$ for $\forall s \in S, \forall v \in V'$, $L$, $N_c$, $C$, $(sc) \in R_c$: the set of connection requests for content $c \in C$, $s \in S$ and $k$: the minimum number of replicas of content.
Output:
    The set of DCN hosting nodes for the content $c$ placement: $A_c, \forall c \in C$.
$A_c = \varnothing, \forall c \in C$;
for each $c \in C$ do
    for each $(sc) \in R_c$ do
        $v$=arg$_{v \in L}$ min$\{PF_{sv} + L_{sv}\}$;
        if $(v \notin A_c)$ then
            $A_c = A_c \bigcup \{v\}$;
        end if
    end for
end for
for each $c \in C$ do
    while $(|A_c| < k)$ do
        $Risk_{min} = \infty$;
        for each $v \in (L - A_c)$ do
            $Risk_v = \sum_{\{(sc) \in R_c, \forall s \in S\}} (PF_{sv} + L_{sv})$;
            if $(Risk_v < Risk_{min})$ then
                $Risk_{min} = Risk_v$; $u = v$;
            end if
        end for
        $A_c = A_c \bigcup \{u\}$;
    end while
    while $(|A_c| > N_c)$ do
        $Risk_{min} = \infty$;
        for each $v \in A_c$ do
            $Risk_v = 0$;
            $Risk_v = \sum_{(sc) \in R_c, \forall s \in S} min_{\forall u \in (A_c - \{v\})} (PF_{su} + L_{su})$
            if $(Risk_v < Risk_{min})$ then
                $Risk_{min} = Risk_v$; $v' = v$;
            end if
        end for
        $A_c = A_c - \{v'\}$;
    end while
end for
return $A_c, \forall c \in C$.
```

(3.13) and (3.14) in Section 3.2.3. First, we find the candidate placement nodes in L for each content $c \in C$, which induce the minimum value of total failure probability of requesting paths and total traffic transmission delay for contents provided by these nodes in lines 2-9. Then, for each content $c \in C$, in lines 11-20 we ensure the number of replicas of content c to satisfy constraint (3.13). In each iteration, a DCN hosting node u in $L - A_c$ is selected as the content c hosting node, which induces the minimum value of total failure probability of requesting paths and total traffic transmission delay for content c when content c is provided by this node. Here, the value of total failure probability and total traffic transmission delay is obtained in line 14. Then, constraint (3.14) is satisfied by the iteration from lines 21-31. In each iteration, a DCN hosting node $v' \in A_c$ is selected, and then removed from A_c. The value of total failure probability of requesting paths and total traffic transmission delay for content c is calculated in line 25 when content c is provided by arbitrary $|A_c| - 1$ nodes in A_c, in which the minimum value is obtained when the node v' is removed from A_c.

Notice that content placement can be optimized either periodically according to daily content requests variation, or within the early warning time of an upcoming disaster if the DCN failure risk is observed higher than the current risk evaluation. For an upcoming disaster with an early warning time, in order to minimize content loss, we need to re-optimize content placement within the early warning time. Since ILP is not scalable in terms of its long running time (i.e., optimal solution may not be found within the given early warning time), time-efficient heuristic is necessary to produce real-time response. On the other hand, solving the ILP for optimal joint design of DCN and content placement with transmission delay optimization is not an easy task for large-scale networks. To this end, we need a time-efficient heuristic as well for scalability.

3.3.2 Complexity Analysis

In this subsection, we analyze the complexity of heuristic for DCN and content placement. The complexity of the Algorithm 3 is dominated by the iterations. The complexity in line 2 is $O(|C|)$. The complexity of the iteration from lines 3-8 is $O(|V'| \times |R_c|)$ and the complexity of the iteration from lines 10-21 is $O((N_d - 1) \times (|V'| - 1) \times |S| \times |C| \times N_d)$. Thus, the total complexity of Algorithm 3 is no more than $O(|C| \times (N_d)^2 \times |V|^2)$.

The complexity of Algorithm 4 is also dominated by the iterations. The complexity of the iteration in lines 2-9 is $O(|C| \times max_{c \in C}(|R_c|) \times |L|)$. The time for the iteration in lines 11-20 is $O((k - min_{c \in C}(|A_c|)) \times (|L| - min_{c \in C}(|A_c|)) \times max_{c \in C}(|R_c|))$. The time for the iteration in lines 21-31 is $O((max_{c \in C}(|A_c|) - N_c) \times max_{c \in C}(|A_c|) \times max_{c \in C}(|R_c|) \times (max_{c \in C}(|A_c|) - 1))$. Thus, the total complexity of Algorithm 4 is no more than $O(|C| \times |N_d|^3 \times |V|)$. From the complexities of these two algorithms, we can find that the complexity of the proposed heuristic is no more than $O(|C| \times |V|^2 \times (N_d)^2)$. Thus, the proposed heuristic runs in polynomial time.

3.4 Numerical Results

In this section, we carry out numerical experiments based on the gridded data of U.S. national seismic hazard map [69]. Assume that the network is deployed in a rectangle area with length 2402 and height 1018. We first demonstrate the proposed vulnerability assessment scheme in Section 3.4.1. Based on the vulnerability information of a given network, we further validate the efficiency of the proposed ILP in Section 3.4.2 and heuristic in Section 3.4.3 for DCN and content placement. For DCN and content placement, Gurobi 6.0 is used to solve the ILP in (3.8)-(3.16). We run the ILP and heuristic algorithms on a computer that has an Intel Core(TM) i3-4030U CPU @ 1.90GHz and 4GB memory and also develop a simulator to emulate

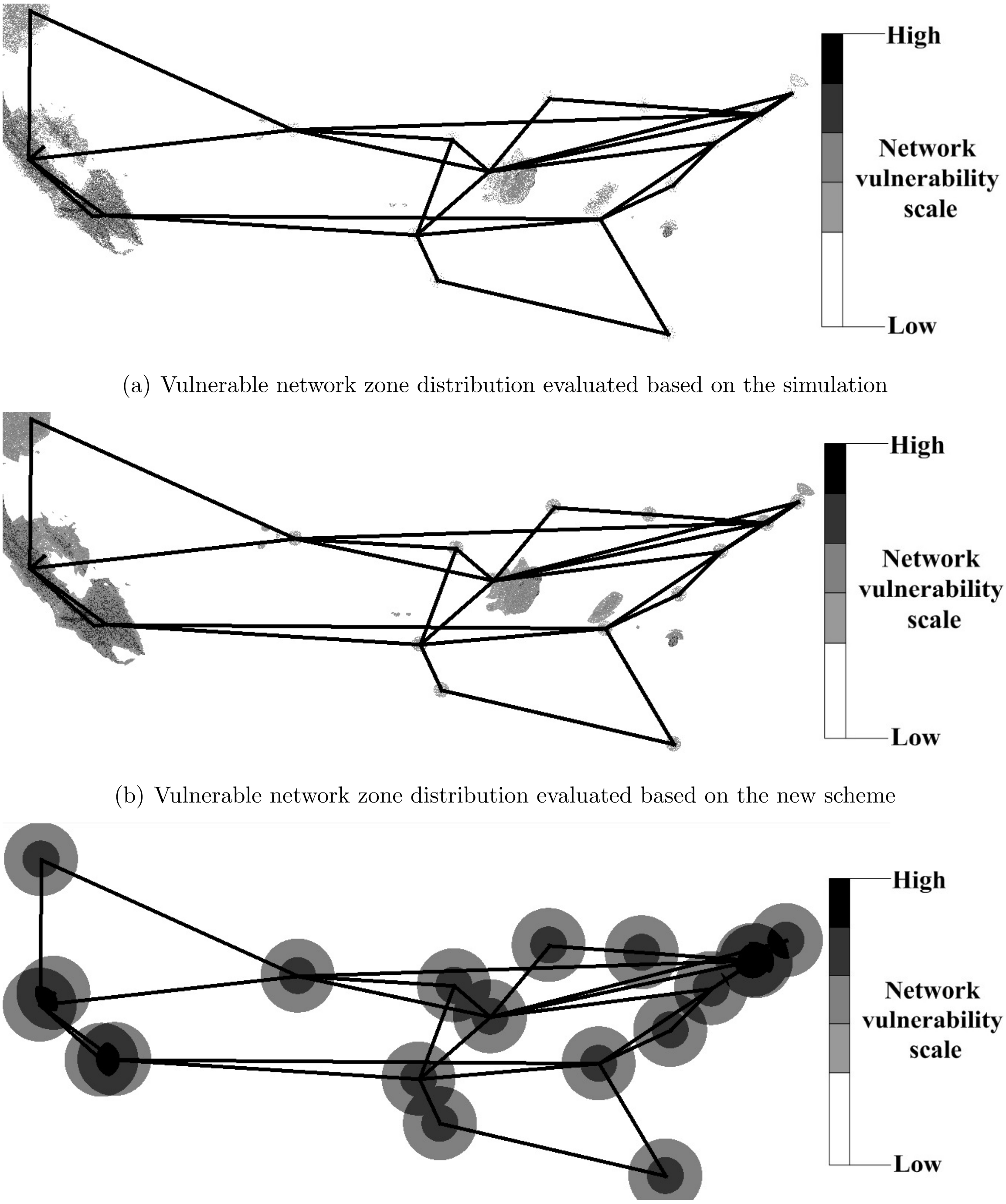
Table 3.4: Parameter settings for PRF model

g from Fig. 1.1	r_1	r_2	p_1	p_2
0.8	100	200	0.95	0.75
0.4	60	120	0.8	0.6
0.3	50	100	0.6	0.3
0.2	25	50	0.5	0.25
others	10	20	0.25	0.1

the random connection requests between nodes and contents. Given a network for DCN and content placement, the simulator generates a random integer of x between 1 and $|C|$ as the number of content requests from each requesting node. The simulator also ensures that each content is requested.

3.4.1 Vulnerability Assessment

For network vulnerability assessment, we consider the U.S. InternetMCI network in Fig. 3.2 with 19 nodes and 33 links, where the length of the shorter segment of link ξ is fixed as 20. To evaluate the vulnerability of network deployed in U.S. due to the non-uniform distribution of a potential earthquake in U.S., we use the grid partition scheme to divide the network area into 1201 × 509 square cells with a side length 2 for each according to the gridded data of U.S. national seismic hazard map. Each PRF is defined by two concentric circles with radiuses (r_1, r_2) and probabilities (p_1, p_2). Since the gridded data of U.S. national seismic hazard map only contains the information of grid partition and peak ground acceleration (g), we can not obtain concrete occurring probability of a PRF falling within one cell from the gridded data of U.S. national seismic hazard map. To facilitate the vulnerability assessment, due to the fact that the gridded data of U.S. national seismic hazard map is obtained based on the map in Fig. 1.1 with an exceedance probability of 2% in 50 years, we set the occurring probability of a PRF falling within one cell as a random value between 0.02 and 0.5. For the PRF with center falling within one cell, we take the


High
Network vulnerability scale
Low
High
Network vulnerability scale
Low
High
Network vulnerability scale
Low

(a) Vulnerable network zone distribution evaluated based on the simulation

(b) Vulnerable network zone distribution evaluated based on the new scheme

(c) Vulnerable network zone distribution evaluated based on the old scheme

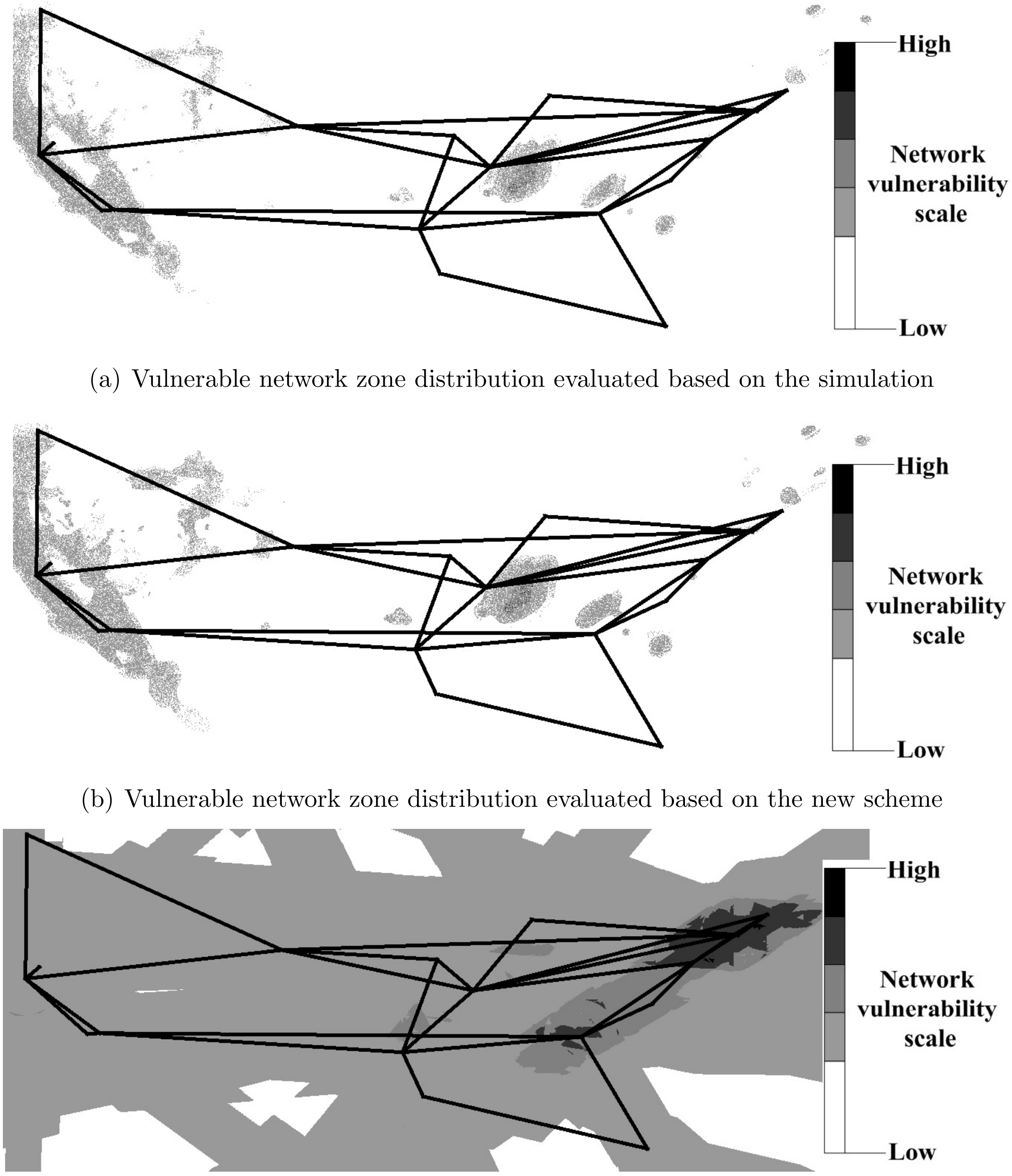
Figure 3.4: Illustration of NFP vulnerable network zone distribution for all nodes

center point of the cell as the center of the PRF and set its parameters r_1, r_2, p_1 and p_2 according to the peak acceleration data (g) in the cell from the gridded data of U.S. national seismic hazard map. The parameter settings are shown in Table 3.4. We compare our vulnerability assessment results with that in [33] and [34] and that based on simulation, respectively. For simplicity, the vulnerability assessment based on our proposed scheme is referred to as new scheme, while that based on [33] and [34] is referred to as old scheme.

In our simulation, we randomly generate a location in each cell as the center of a PRF when the PRF occurs in this cell. Other parameter settings keep the same as above. Here, we have carried out 10 different simulations, and then the vulnerability for each node (or link) is evaluated by the average of all simulation results. For vulnerability assessment in [33] and [34], the parameters r_1, r_2, p_1, and p_2 of the PRF are the same and fixed as $r_1 = 50$, $r_2 = 100$, $p_1 = 0.60$, and $p_2 = 0.30$, and the occurring probability of a PRF falling within one cell is uniformly distributed.

Fig. 3.4 illustrates the NFP vulnerable network zone distributions for all nodes under the simulation, the new scheme and the old scheme, respectively. The results in Fig. 3.4 clearly indicate that the NFP vulnerable network zone distribution for all nodes based on the new scheme generally complies with the simulation results and both of them match the potential earthquake distribution in U.S. as illustrated in Fig. 1.1. These results show that our proposed vulnerability assessment scheme is efficient to evaluate the vulnerability of nodes due to the real disaster. It is notable that since the old scheme does not take the global non-uniform distribution of a disaster in terms of its occurring probability and intensity into account, the NFP vulnerable network zone distribution for all nodes based on the old scheme is quite different from that based on the new scheme.

Fig. 3.5 shows the LFP vulnerable network zone distribution for all links, with Fig. 3.5(a) for the simulation, Fig. 3.5(b) for the new scheme and Fig. 3.5(c) for the



(a) Vulnerable network zone distribution evaluated based on the simulation

(b) Vulnerable network zone distribution evaluated based on the new scheme

(c) Vulnerable network zone distribution evaluated based on the old scheme

Figure 3.5: Illustration of LFP vulnerable network zone distribution for all links

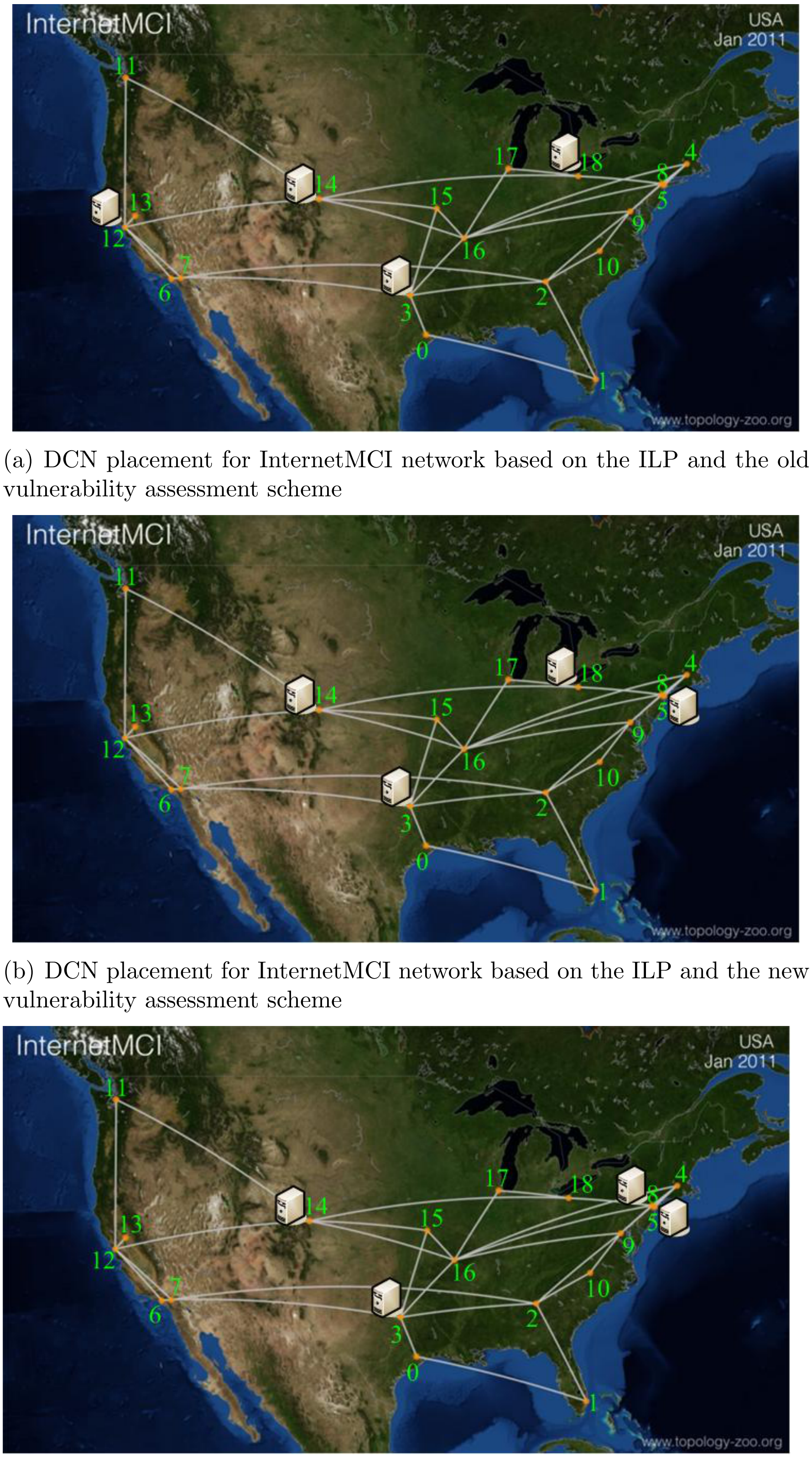
old scheme, respectively. From Fig. 3.5 we can get similar conclusions as those in Fig. 3.4. We can also observe that our proposed scheme is efficient to evaluate the vulnerability of links due to a region failure. Based on such vulnerable network zone distribution for all nodes or links, we can easily identify the most vulnerable network zones, i.e., the zones in which the PRF falling within each cell has the most significant impact to the network nodes or links. Since our proposed vulnerability assessment can efficiently evaluate the impact to a network from a real disaster, "vulnerability map" based on the new scheme will be very helpful for us to identify the optimal DCN and content placement in a given network against the disaster.

In our experiments, the old scheme only takes uniformly distributed data, but a real disaster generally entails the non-uniform case. The new scheme considers non-uniform distribution and it covers the old scheme as a special case. By using uniform distribution for the old scheme in our experiments, we indeed intend to show the drawback of vulnerability assessment under uniform distribution, rather than comparing with the new scheme.

3.4.2 Placement in Small-Scale Networks

For our proposed ILP framework for DCN and content placement, we also consider the U.S. InternetMCI network with 19 nodes and 33 links. In our simulation, we set $\beta = 100$ and $\delta = 1000$ (i.e., the second DCN and content placement scenario in Section 3.4.4). We consider 4 DCNs and 20 contents, and each content has at least 2 and at most 3 replicas. All nodes in the network are set as candidate placement nodes for DCNs. The "vulnerability map" for such network is obtained by the vulnerability assessment with the same parameters as in Section 3.4.1. To facilitate the calculation, we convert the values of L_{sv} to values between 0 and 1.

The DCN placement based on the ILP and the old vulnerability assessment scheme is shown in Fig. 3.6(a) and that based on the ILP and the new vulnerability assess-



(a) DCN placement for InternetMCI network based on the ILP and the old vulnerability assessment scheme

(b) DCN placement for InternetMCI network based on the ILP and the new vulnerability assessment scheme

(c) DCN placement for InternetMCI network based on the heuristic and the new vulnerability assessment scheme

Figure 3.6: DCN placement scenarios

Table 3.5: Content placement in DCNs for ILP and heuristic

ILP		Heuristic	
DCNs	Contents	DCNs	Contents
3	0,1,3,4,5,8,10,11 12,13,14,16,17	3	0,1,3,4,5,8,10,11,12 13,14,15,16,17
5	0,3,4,5,7,8,9,11 12,13,15,19	5	0,3,4,5,7,8,9,11 12,13,15,19
14	1,2,3,4,5,6,7,8,9,10,11 13,14,15,16,17,18,19	8	0,1,2,6,7,8,10,13 16,17,18,19
18	0,1,2,6,7,9,12 15,16,18,19	14	1,2,3,4,5,6,7,9,10,11,12 14,15,16,17,18,19

ment scheme is shown in Fig. 3.6(b). From Figs. 3.6(a) and (b) we can find that DCNs are placed at nodes 3, 12, 14 and 18 for the former and 3, 5, 14 and 18 for the latter. Besides, the failure risk is 97.49 (calculated based on the new vulnerability information) for the former and 96.15 for the latter. Although the gap of failure risk is only 1.39% (i.e., $(97.49 - 96.15)/96.15$) between the above two scenarios, the total failure probability of DCN hosting nodes can be dramatically reduced under the new vulnerability assessment scheme (the total failure probability of DCN hosting nodes is 0.008111 for the former and 0.000438 for the latter).

Fig. 3.6(c) shows the DCN placement based on the heuristic solution and the new vulnerability assessment scheme. The DCNs are placed at nodes 3, 5, 8 and 14 and the failure risk is 100.21. Thus, the gap of failure risk between the ILP and heuristic under the new vulnerability assessment scheme is 4.22% (i.e., $(100.21-96.15)/96.15$). The contents hosted at each DCN are shown in Table 3.5 for ILP and heuristic under the new vulnerability assessment scheme, respectively. Table 3.5 shows that the same DCN hosting node determined by the ILP and heuristic contains similar contents. From Fig. 3.3 and 3.4, we can also find that the DCN hosting nodes based on the ILP and heuristic under the new vulnerability scheme avoid the nodes with high NFP and the most vulnerable network zones for all nodes. Besides, under the new vulnerability assessment scheme we also have carried out other experiments for 10 different groups of connection requests generated randomly with similar network size. The average gap of failure risk between the ILP and heuristic is 3.5%, which confirms the superior performance of the proposed heuristic.

3.4.3 Placement in Large-Scale Networks

To verify the performance of our proposed heuristic for large-scale networks under the new vulnerability assessment scheme, we randomly generate a network by simulator with 100 nodes and 202 links. In order to reduce the complexity, in this experiment we only consider link-disjoint k-shortest paths between an arbitrary pair of nodes to implement routes (k=3). The "vulnerability map" for this network is obtained in a similar way as that of U.S. InternetMCI network. Except the number of DCNs and contents to be placed, other parameter settings are similar to those in Section 3.4.2.

The performance of ILP and heuristic for the cases $|C|$={10, 20, 30, 40, 50} are summarized in Table 3.6 when the number of DCNs to be placed is 4. In Table 3.7, we show the performance of ILP and heuristic for the cases N_d={4, 8, 12, 16, 20, 24, 28, 32, 36, 40} when the number of contents to be placed is 10. From Tables 3.6 and 3.7, we can observe that our proposed heuristic is more scalable, and the ILP is sensitive to $|C|$. Table 3.7 also shows that the running time of ILP decreases and that of heuristic increases when N_d increases, but the running time of heuristic increases slowly. Besides, from Tables 3.6 and 3.7 we can find that although the gaps of failure risk between the ILP and heuristic vary with the increases of $|C|$ and N_d, their sensitivities to the variations of $|C|$ and N_d are different. For a fixed number of DCNs to be placed of N_d=4 and when we increase the number of contents to be placed $|C|$ from 10 to 50, the average gap of failure risk is 11.2%. When we increase the number of DCNs to be placed N_d from 4 to 40 at a fixed number of contents to be placed of $|C|$=10, the average gap of failure risk is 26.4%.

Table 3.6: Performance analysis in large-scale network with 4 DCNs and different numbers of contents for ILP and heuristic

$\lvert C\rvert$	ILP		Heuristic	
	Failure risk	Running time (seconds)	Failure risk	Running time (seconds)
10	60.74	16.64	68.86	0.44
20	115.43	181.61	128.55	0.51
30	163.03	242.99	185.96	0.53
40	210.15	476.88	228.55	0.56
50	264.74	3297.55	286.57	0.59

Table 3.7: Performance analysis in large-scale network with 10 contents and different numbers of DCNs for ILP and heuristic

N_d	ILP		Heuristic	
	Failure risk	Running time (seconds)	Failure risk	Running time (seconds)
4	60.74	16.64	68.86	0.44
8	60.45	12.11	74.66	0.56
12	60.67	11.86	76.76	0.61
16	60.94	11.83	79.85	0.71
20	61.21	12.62	80.32	0.8
24	61.49	11.95	78.24	0.88
28	61.77	12.02	80.05	1.04
32	62.05	12.38	79.66	1.039
36	62.34	10.39	80	1.22
40	62.62	10.14	78.22	1.49

3.4.4 Effect of Scaling Factor δ on Placement

The scaling factor δ is used to control the weight among the total failure probability of DCN hosting nodes, the total failure probability of requesting paths and the total traffic transmission delay. Thus, for different values of δ, we can obtain different DCN and content placement scenarios. Considering the ILP with same simulation settings in Section 3.4.2, for different values of δ, there are five different DCN and content placement scenarios. In Table. 3.8, we show the total failure probability of DCN hosting nodes $\sum\limits_{v\in V'} H_v PF_v$ (abbreviated as DFP) and the total failure probabil-

Table 3.8: Tradeoff between DFP and $PFP + TD$

Placement Scenarios	DFP	PFP	TD
1	0.008111	19.31	70.07
2	0.000438	20.73	74.98
3	0.000329	22.31	79.69
4	0.000283	23.7	84.29
5	0.000282	24.54	87.26

ity of requesting paths $\sum_{v \in V'} \sum_{s \in S} \sum_{c \in C} H_v^{sc} PF_{sv}$ (abbreviated as PFP) as well as the total traffic transmission delay $\sum_{v \in V'} \sum_{s \in S} \sum_{c \in C} H_v^{sc} L_{sv}$ (abbreviated as TD) for five DCN and content placement scenarios, respectively. From Table. 3.8, we can find a desirable tradeoff between DFP and $PFP + TD$ by adjusting the value of δ in the DCN design phase.

3.5 Summary

We studied the DCN and content placement problem under global non-uniform distribution of a potential region failure due to disaster in large-scale geographical areas. By proposing a general grid partition-based vulnerability assessment scheme, we can determine the "vulnerability map" of a given network for DCN and content placement, which provides an important input for our proposed ILP and heuristic. Based on the vulnerability map, our proposed ILP can generate optimal DCN and content placement solutions to minimize the DCN failure risk due to disaster. This achieves best-effort protection of DCN and content against the region failure. To make our solution more scalable for large-scale networks, a heuristic was further proposed. Numerical results showed that our work can lead to a more feasible solution. It can well protect DCN and content under global non-uniform distribution of the potential region failure scenario.

CHAPTER IV

Homogeneous Data Backup Based on Early Warning of Region Failure

In this chapter, we study the homogeneous data backup based on early warning of region failure. We assume that there is only one data center network (DCN) node falling within the region that will be affected by a disaster after ε early warning time (referred to as threatened DCN node hereafter). We consider urgent backup within the early warning time ε where different types of data at the threatened DCN node are backed up to the same set of backup DCN nodes. To this end, we divide our design into two sub-problems: Backup Capacity Evaluation (BCE) and Backup Cost Minimization (BCM). The former helps DCN operators to find the maximum backup capacity, and thus fully utilize the ε early warning time to back up as much data as possible. Since the maximum backup capacity may not be sufficient for backing up all data, priority can be given to those more important data. The latter minimizes backup cost by properly selecting a set of safe backup DCN nodes and routes for those more important data. Both integer linear programs (ILPs) and heuristic are proposed for the two sub-problems. Extensive numerical results show that the proposed algorithms can automatically adapt to different early warning times ε for generating cost-efficient data backup solutions.

Ii is notable that in this chapter, data can be stored in either a distributed manner

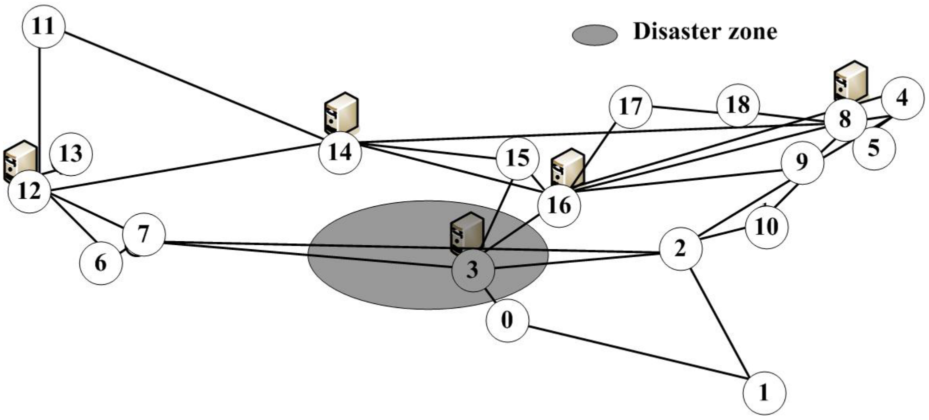


Figure 4.1: Geo-distributed DCNs with DCN node 3 threatened by a disaster and other DCN nodes serving as candidate backup nodes.

(across multiple DCNs) or a centralized manner (at a single DCN). Multiple replicas at different DCNs are allowed as well. Generally, a DCN backs up data or replicas periodically in its normal operation state when no disaster presents. At the earliest time when the DCN is aware of the disaster, some data may not have been successfully backed up in the current period, or still in an unsynchronized state. To this end, the work in this chapter considers to protect such unsafe data.

4.1 Network Model

We assume multiple DCNs in an optical backbone network, each hosted by a distinct node. We also assume that data is transmitted in the network through all-optical paths where the OXCs (optical cross-connects) with wavelength converters (i.e., wavelength conversion capabilities) are used at intermediate nodes for transparent optical connections. Network topology is denoted by a graph $G(V, E)$, where V is the set of all nodes and E is the set of all fiber links. There is a single threatened DCN node that will be affected by a disaster after ε early warning time. Other DCN nodes will not be affected by the disaster, and they can serve as candidate backup DCN nodes. Each candidate backup DCN node has a certain amount of backup storage, whereas online bandwidth available on each link at the disaster time can be measured by the DCN operator. Fig. 4.1 gives an example of the U.S. InternetMCI network [70] with five geo-distributed DCNs hosted at nodes 3, 8, 12, 14 and 16. Suppose DCN node 3 will be affected by a disaster after ε early warning time, as shown by the shaded area in Fig. 4.1. Data at the threatened DCN node 3 can be backed up to the backup DCN nodes 8, 12, 14, and 16. Backup cost consists of data storage and transmission costs. The former is the sum of costs of required storage (counted in data units) at all backup DCN nodes, where W_v denotes the storage cost per data unit at backup DCN node v. The latter counts for the costs of working wavelength capacity (including the costs of necessary wavelength converters) in all backup routing paths.

In general, a DCN service provider (such as Google) needs to consider the disaster scenario at the network planning stage. As a result, multiple DCNs are deployed in different geographical regions to avoid simultaneous failures [4]. Nowadays, such a geo-distributed DCN architecture is well supported by long-haul optical inter-connects under the WDM (Wavelength Division Multiplexing) technology. In this thesis, we assume that multiple DCNs are affiliated to a single DCN provider and there is only one threatened DCN node.

Note that store-and-forward schemes (using safe DCNs as relays to forward data) are not considered in this thesis. This is because we assume only a single threatened DCN, and other DCNs (including those may possibly serve as intermediate DCNs in a store-and-forward scheme) are taken as safe DCNs. As a result, data will be safely protected as long as they can arrive at such a safe DCN.

4.2 ILP Formulations

In this section, we first provide an ILP to solve BCE under the ε early warning time constraint, which can be used to determine the amount of data that should be backed up in the threatened DCN node. For the determined amount of data that should be backed up, we also develop another ILP to solve BCM by identifying the optimal selections of backup DCN nodes and routes, such that the overall data backup cost is minimized.

4.2.1 Notation List

The detailed inputs and variables used in the ILP formulations are listed in Tables 4.1 and 4.2.

Table 4.1: Parameters for Inputs

Notation	Definition
V	Similar to the definition in Section 3.2.2.
E	Similar to the definition in Section 3.2.2.
$\bar{V} \subset V$	The set of all backup DCN nodes in network $G(V, E)$.
ε	The early warning time of disaster for backing up data (It is quantified with the number of time units).
$P = \{p \mid p =< Sc, De_p, L_p >\}$	The set of paths between the threatened DCN node and the backup DCN nodes where Sc, De_p, L_p are source DCN node (i.e., threatened DCN node), destination DCN node (i.e., backup DCN node), and the set of links on path p.
$D =< Sc, VL, Vl >$	The data in the threatened DCN node where Sc is the threatened DCN node, VL is the amount of data D and Vl is the amount of data that can be backed up, i.e., $Vl \leq VL$ (VL and Vl are quantified with the number of data units).
Re	The transmission rate of each wavelength (It is quantified with the number of data units that are transmitted by one wavelength per time unit).

S_v	The available storage capacity in DCN node $v \in \bar{V}$ (It is quantified with the number of data units).
B_e	The available bandwidth on link $e \in E$ (It is counted in the number of wavelength channels).
W_v	The cost of a data unit stored in the DCN node $v \in \bar{V}$
W_e	The cost of a wavelength on link $e \in E$.
$A_p^e \in \{0, 1\}$	It equals to 1 if link $e \in L_p, p \in P$.
λ	Predefined constant larger than $max\{B_p, Su_v \mid \forall v \in \bar{V}, \forall p \in P\}$.

Table 4.2: Variables

Notation	Definition
U_v	Binary variable. It takes 1 if the DCN node $v \in \bar{V}$ is used for backing up data and 0 otherwise.
U_p	Binary variable. It takes 1 if the path $p \in P$ is used for backing up data and 0 otherwise.
Su_v	Non-negative integer. It is the used storage capacity in node $v \in \bar{V}$ for backing up data.
B_p	Non-negative integer. It is the used bandwidth on path $p \in P$ for backing up data.
M_ε	Non-negative integer. It is the total amount of data that can be backed up in the threatened DCN node within time ε.

4.2.2 ILP for Backup Capacity Evaluation

$$Maximize\Big\{M_{\varepsilon}\Big\}. \tag{4.1}$$

Subject to

$$Su_v \leq S_v, \forall v \in \bar{V}; \tag{4.2}$$

$$\sum_{v \in \bar{V}} Su_v = M_{\varepsilon}; \tag{4.3}$$

$$\sum_{p \in P} A_p^e B_p \leq B_e, \forall e \in E; \tag{4.4}$$

$$M_{\varepsilon} \leq \sum_{v \in \bar{V}} S_v; \tag{4.5}$$

$$\sum_{v \in \bar{V}} U_v \geq 1; \tag{4.6}$$

$$U_p \leq \frac{U_{De_p} + 1}{2}, \forall p \in P; \tag{4.7}$$

$$\sum_{p \in P, De_p = v} U_p \geq U_v, \forall v \in \bar{V}; \tag{4.8}$$

$$U_p \leq B_p, \forall p \in P; \tag{4.9}$$

$$U_p \geq B_p/\lambda, \forall p \in P; \tag{4.10}$$

$$U_v \leq Su_v, \forall v \in \bar{V}; \tag{4.11}$$

$$U_v \geq Su_v/\lambda, \forall v \in \bar{V}; \tag{4.12}$$

$$\frac{Su_v}{\sum\limits_{p \in P, De_p = v} B_p} \leq \varepsilon \cdot Re, \forall v \in \bar{V}. \tag{4.13}$$

Objective (4.1) maximizes the total amount of data that can be backed up. Constraint (4.2) ensures that the used storage capacity in a backup DCN node for backing up data does not exceed the available storage capacity of this DCN node. Constraint (4.3) guarantees that data with the amount M_ε can be backed up to the backup DCN nodes. Constraint (4.4) ensures that the used bandwidth for backing up data on a link does not exceed the available capacity of this link. Constraint (4.5) guarantees that the amount of data that can be backed up does not exceed the total available storage capacity of all backup DCN nodes. Constraint (4.6) guarantees that data is backed up to at least one backup DCN node. Constraint (4.7) implies that if a path is selected for backing up data, then the destination node of this path must be selected as the backup DCN node for storing data. Constraint (4.8) implies that if a DCN node is selected as the backup node for storing data, then at least one path destined to such DCN node must be selected as the transmission path for backing up data. Constraints (4.9) and (4.10) define U_p while constraints (4.11) and (4.12) define U_v. Here, we use λ larger than $max\{B_p, Su_v | \forall v \in \bar{V}, \forall p \in P\}$ to ensure that the constraints (4.10) and (4.12) can be properly established when $U_p = 1, B_p > 0$ and $U_v = 1, Su_v > 0$, respectively. Constraint (4.13) ensures that the time for backing up data does not exceed the time ε.

4.2.3 ILP for Backup Cost Minimization

After we achieve the maximum amount of data M_ε^u that can be backed up within the given ε based on the above ILP, we can determine the amount of data D that should be backed up. For the determined amount of data D that should be backed up, we then develop an ILP shown as follows to generate optimal solutions of backup DCN nodes and routes under the time ε constraint, such that the overall data backup cost is minimized.

$$Minimize\Big\{\sum_{v\in\bar{V}} W_v Su_v + \sum_{p\in P}\sum_{e\in L_p} W_e B_p\Big\}. \tag{4.14}$$

Objective (4.14) minimizes the overall data backup cost, which consists of two terms. The first term is the cost of storing all data that should be backed up and the second term is the total bandwidth cost for transmitting the data that should be backed up. The constraints in such ILP are similar as those in Subsection 4.2.2 in which M_ε is replaced by the amount of determined data that should be backed up $Vl, (Vl \leq min(M_\varepsilon^u, VL))$ in the constraints (4.3) and (4.5). Although the ILP for

BCM has two terms, they can be integrated into a single objective for cost minimization, with storage and transmission costs counted into a total backup cost. Therefore, BCM can be taken as a single-objective optimization. Since we assume wavelength converters at intermediate nodes (if necessary) in the network for transparent optical connections, our ILP models can simply count the bandwidth (i.e., the number of available wavelengths) on each link to ensure non-overlapping wavelengths.

To simplify our analysis, we assume static network status within the early warning time. Nevertheless, this assumption can easily be extended to the scenario where network status changes within the early warning time. In this case, we can divide the early warning time into multiple time intervals within which network status can be taken as static for each. Then, our ILPs can be applied in each time interval for data backup. Similar technique is also used in [44].

4.3 Heuristic

Since solving ILP for large-scale problems (e.g. a large amount of data to be backed up and a large number of backup DCN nodes deployed in a large-scale network) is intractable, it is generally hard to get an optimal ILP solution for data backup in real-time. To make our approach more scalable, in this section we propose a time-efficient heuristic for BCE and BCM to meet the practical engineering requirement. It is notable that to solve BCE by the heuristic, we only need to set the amount of data to be backed up (Vl) as the total available capacity of all backup DCN nodes (i.e., $Vl = \sum_{v \in \bar{V}} S_v$), and then the amount of data (Vl) to be backed up for BCM is determined according to the result from BCE where $Vl \leq min(M_{\varepsilon}^{u}, VL)$.

4.3.1 Algorithm Description

The proposed heuristic is illustrated in Algorithm 1 which includes two procedures, i.e., Integer Data Backup and Remainder Data Backup. Here, $\bar{V}$, Re, D, P, S_v,

```
Algorithm 1 Data Backup (DBu):
Input:
    $G(V,E)$, $\bar{V} \subset V$, $Re$, $D$, $P$, $S_v$ and $W_v$ for $\forall v \in \bar{V}$, $B_e$ and $W_e$ for $\forall e \in E$, and
    the time for backing up data $\varepsilon$.
Output:
    The backup scheme, i.e., the sets of backup DCN nodes ($V_b$) and backup trans-
    mission paths ($T_p$) for data $D$, the overall backup cost $Cost$ and the total amount
    of data that can be backed up within time $\varepsilon$, $M_\varepsilon$.
1: Set $V_b = \varnothing$, $T_p = \varnothing$, $Cost = 0$, $M_\varepsilon = 0$;
2: Set $S_v^I = \lfloor \frac{S_v}{\varepsilon \cdot Re} \rfloor \cdot \varepsilon \cdot Re$ for $\forall v \in \bar{V}$, $Vl_I = \lfloor \frac{Vl}{\varepsilon \cdot Re} \rfloor \cdot \varepsilon \cdot Re$ for data $D$;
3: Set $S_v^R = S_v - S_v^I$ for $\forall v \in \bar{V}$, $Vl_R = Vl - Vl_I$ for data $D$;
4: Call Procedure Integer Data Backup;
5: Set $Vl_R = Vl_R + Vl_I$ for data $D$;
6: Set $S_v^R = S_v^R + S_v^I$ for $\forall v \in \bar{V}$;
7: Call Procedure Remainder Data Backup.
```

W_v, B_e, W_e, ε and M_ε are defined in Section 4.2.1, and let $|A|$ denote the number of elements in an arbitrarily given set A. Note that in the proposed heuristic the bandwidth on transmission path is assigned with the integer.

In Algorithm 1, the initialization is first shown in lines 1-3 where we set $V_b = \varnothing$, $T_p = \varnothing$ for data D, $Cost = 0$ and $M_\varepsilon = 0$, and the available capacity of each backup DCN node and the amount of data D that should be backed up are divided into two parts, respectively, i.e., S_v^I and S_v^R for $\forall v \in \bar{V}$, Vl_I and Vl_R for data D. Then we call Procedure 1 (i.e., Integer Data Backup) by taking S_v^I, Vl_I and the inputs of Algorithm 1 as its inputs. After executing Procedure 1, Procedure 2 (i.e., Remainder Data Backup) is executed by taking S_v^R, Vl_R and the inputs of Algorithm 1 that updated by Procedure 1 as its inputs.

Integer Data Backup: In this procedure, for data D with the amount Vl that should be backed up, we consider only to back up the amount of data Vl_I. In line 2, for BCE, we select a path p ($S_{D_{e_p}}^I > 0$) from the set P which has nonzero available bandwidth and the ability to back up the largest amount of data, where the available bandwidth on path p is $Min_{e \in p}\{B_e\}$ and the amount of data that can be backed up

Procedure 1 Integer Data Backup (IDBu):

```
1: while ($Vl_I > 0$) do
2:   Select a path $p$, ($S^I_{De_p} > 0$) with nonzero available bandwidth $Min_{e \in p}\{B_e\}$ and the ability to back up the largest amount of data based on (4.15) from the set $P$ for BCE (Select a path $p$, ($S^I_{De_p} > 0$) with nonzero available bandwidth $Min_{e \in p}\{B_e\}$ and the smallest cost based on (4.16) from the set $P$ for BCM);
3:   if ($p$ is found) then
4:     Determine a bandwidth $B_p = Min(\frac{S^I_{De_p}}{\varepsilon \cdot Re}, \frac{Vl_I}{\varepsilon \cdot Re}, Min_{e \in p}\{B_e\})$ on path $p$;
5:     Set $Vl_I = Vl_I - B_p \cdot \varepsilon \cdot Re$, $S^I_{De_p} = S^I_{De_p} - B_p \cdot \varepsilon \cdot Re$;
6:     Set $B_e = B_e - B_p$ for $\forall e \in p$;
7:     Set $V_b = V_b \bigcup De_p$, $T_p = T_p \bigcup p$;
8:     Set $Cost = Cost + W_{De_p} \cdot B_p \cdot \varepsilon \cdot Re + \sum_{e \in p} W_e \cdot B_p$;
9:     Set $M_\varepsilon = M_\varepsilon + B_p \cdot \varepsilon \cdot Re$;
10:  else
11:    Exit procedure;
12:  end if
13: end while
```

by path p is determined as

$$Min(\frac{S^I_{De_p}}{\varepsilon \cdot Re}, \frac{Vl_I}{\varepsilon \cdot Re}, Min_{e \in p}\{B_e\}) \cdot \varepsilon \cdot Re. \tag{4.15}$$

For BCM, we select a path p ($S^I_{De_p} > 0$) from the set P which has nonzero available bandwidth and the smallest cost. Here, the available bandwidth on path p is $Min_{e \in p}\{B_e\}$ and the cost is determined as

$$\frac{\sum_{e \in p} W_e + W_{De_p} \cdot \varepsilon \cdot Re \cdot Min(\frac{S^I_{De_p}}{\varepsilon \cdot Re}, \frac{Vl_I}{\varepsilon \cdot Re}, Min_{e \in p}\{B_e\})}{\varepsilon \cdot Re \cdot Min(\frac{S^I_{De_p}}{\varepsilon \cdot Re}, \frac{Vl_I}{\varepsilon \cdot Re}, Min_{e \in p}\{B_e\})}. \tag{4.16}$$

If we find an available path p, in line 4, the assigned bandwidth B_p on path p for backing up data D is determined as

$$Min(\frac{S^I_{De_p}}{\varepsilon \cdot Re}, \frac{Vl_I}{\varepsilon \cdot Re}, Min_{e \in p}\{B_e\}). \tag{4.17}$$

Procedure 2 Remainder Data Backup (RDBu):

```
1: while ($Vl_R > 0$) do
2:    Select a path $p$, ($S^R_{De_p} > 0$) with nonzero available bandwidth $Min_{e \in p}\{B_e\}$ and
      the ability to back up the largest amount of data based on (4.15) from the
      set $P$ for BCE (Select a path $p$, ($S^R_{De_p} > 0$) with nonzero available bandwidth
      $Min_{e \in p}\{B_e\}$ and the smallest cost based on (4.16) from the set $P$ for BCM);
3:    if ($p$ is found) then
4:       Determine a bandwidth $B_p = 1$ on path $p$;
5:       if ($S^R_{De_p} \geq Vl_R$) then
6:          Set $Cost = Cost + W_{De_p} \cdot Vl_R + \sum_{e \in p} W_e \cdot B_p$;
7:          Set $M_\varepsilon = M_\varepsilon + Vl_R$;
8:          Set $S^R_{De_p} = S^R_{De_p} - Vl_R$, $Vl_R = 0$;
9:       else
10:         Set $Cost = Cost + W_{De_p} \cdot S^R_{De_p} + \sum_{e \in p} W_e \cdot B_p$;
11:         Set $M_\varepsilon = M_\varepsilon + S^R_{De_p}$;
12:         Set $Vl_R = Vl_R - S^R_{De_p}$, $S^R_{De_p} = 0$;
13:      end if
14:      Set $B_e = B_e - B_p$ for $\forall e \in p$;
15:      Set $V_b = V_d \bigcup De_p$, $T_p = T_p \bigcup p$;
16:   else
17:      Exit procedure;
18:   end if
19: end while
```

The above expression (4.17) ensures that the assigned bandwidth on path p for backing up data satisfies the constraints of the available capacity of DCN node De_p, the amount of data that should be backed up and the available bandwidth on path p. In lines 5-6, we update the values of Vl_I, $S^I_{De_p}$ and B_e for each $e \in p$, respectively. Node De_p is added into set V_b and the path p is also added into set T_p in line 7. The backup cost and the total amount of data that can be backed up are obtained in lines 8-9, respectively. If we can not find an available path p, procedure exits in line 11.

Remainder Data Backup: In this procedure, for data D with the amount Vl that should be backed up, we consider to back up the amount of data Vl_R. In line 2, the path p is selected in the same way as that in Procedure 1. If we find an available path p, we take a ceiling function of $Min(\frac{Vl_R}{\varepsilon \cdot Re}, \frac{S^R_{De_p}}{\varepsilon \cdot Re})$ where the value of

$Min(\frac{Vl_R}{\varepsilon \cdot Re}, \frac{S^R_{De_p}}{\varepsilon \cdot Re})$ is less than 1. Then the assigned bandwidth B_p on path p for backing up data D is set as 1. From lines 5-13, we update the values of Vl_R, $S^R_{De_p}$, $Cost$ and M_ε for two cases (i.e., $S^R_{De_p} \geq Vl_R$ and $S^R_{De_p} < Vl_R$), respectively. In lines 14-15, the value of B_e for $\forall e \in p$, V_b and T_p are updated, respectively. If we can not find an available path p, procedure exits in line 17.

4.3.2 Complexity Analysis

In this subsection, we analyze time complexity of the proposed heuristic. Before calculating the complexity of Algorithm 1, we first give the complexity of Procedure 1. In Procedure 1, for backing up data D with the amount Vl_I, the iteration from lines 1-13 is executed at most $|P|$ times, i.e., we traverse all paths in set P for backing up data. For line 2, since we need to traverse all available paths for backing up data D in set P, the complexity of this operation is no more than $O(|P| \times |E|)$. Besides, the complexity of the operations from lines 4-9 is $O(|E|)$. Thus, the complexity of Procedure 1 is no more than $O(|P|^2 \times |E|)$. From Procedure 2, we can find that it has the same complexity of Procedure 1. Since both of the complexities of the operation from lines 1-3 and that from lines 5-6 in the Algorithm 1 are $O(|\bar{V}|)$, the complexity of the Algorithm 1 is $O(|\bar{V}| + |P|^2 \times |E|)$ and then the proposed heuristic runs in polynomial time.

4.4 Numerical Results

In this section, we carry out numerical experiments on U.S. InternetMCI network with 19 nodes and 33 links to validate the proposed ILP models and heuristic. We assume that there is an ε-time early warning disaster which will affect DCN node 3 after ε time (i.e., DCN node 3 is the threatened DCN node). The number of available wavelength channels (i.e., available bandwidth) on each link is set as a random integer

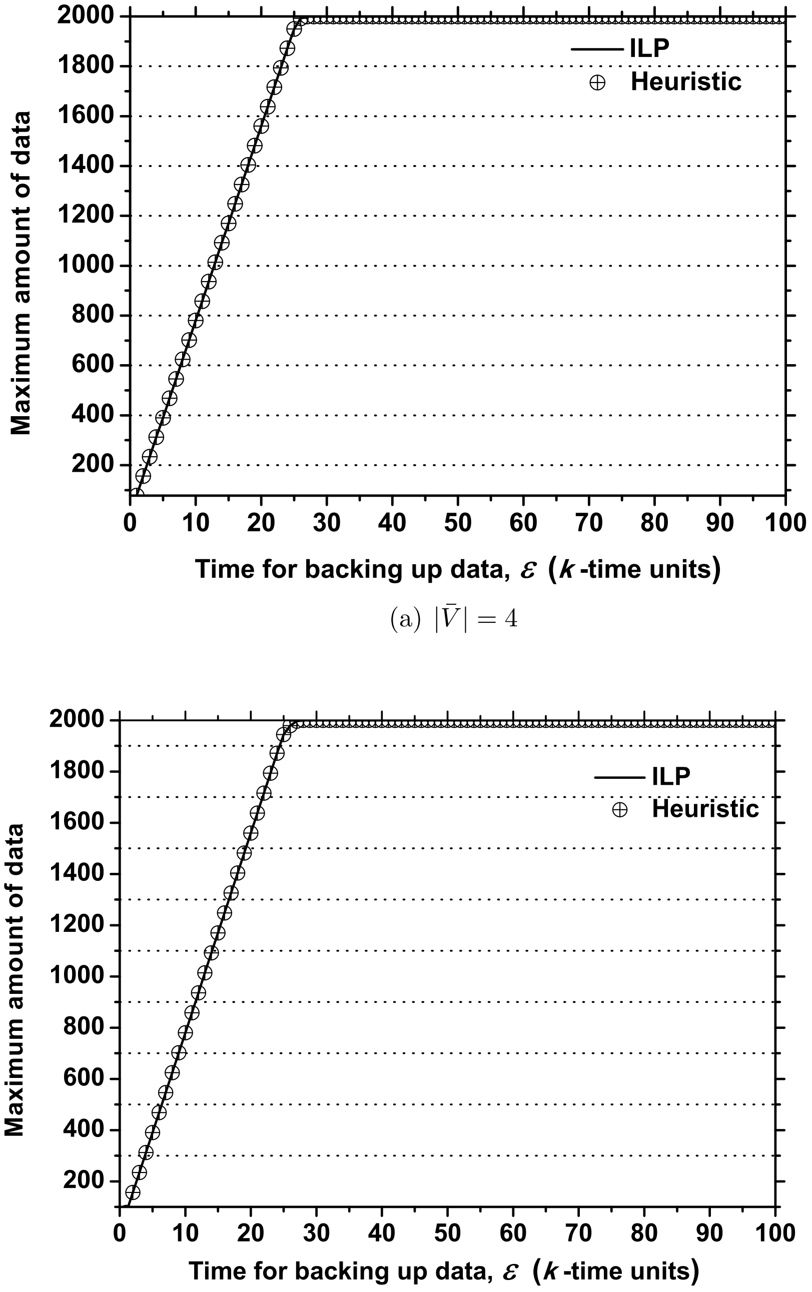
Table 4.3: The costs of a wavelength on each link in the U.S. InternetMCI network

Link	Cost	Link	Cost	Link	Cost
(0,1)	625	(4,8)	105	(9,10)	157
(0,3)	133	(4,9)	240	(9,16)	602
(1,2)	352	(4,16)	826	(11,12)	393
(2,3)	488	(5,8)	9	(11,14)	761
(2,7)	1309	(6,7)	35	(12,13)	49
(2,9)	365	(6,12)	223	(12,14)	701
(2,10)	213	(7,12)	249	(14,15)	423
(3,7)	824	(8,9)	135	(14,16)	532
(3,15)	269	(8,14)	1230	(15,16)	128
(3,16)	256	(8,16)	725	(16,17)	249
(4,5)	99	(8,18)	300	(17,18)	252

between 10 and 30. The total available storage capacity in all backup DCN nodes is set as 2000 data units.

In our experiments, we set the cost of a wavelength on a link as the length of the link. In particular, wavelength cost on each link in the U.S. InternetMCI network is shown in Table 4.3. We set the cost of a data unit stored in backup DCN node as a random value between 40 and 80. We also set $\lambda = 2000$ and $Re = 1$. We here consider two scenarios, i.e., $|\bar{V}| = 4$ backup DCN nodes (i.e., backup DCNs host at nodes 8, 12, 14 and 16) and $|\bar{V}| = 10$ backup DCN nodes (i.e., backup DCNs host at nodes 2, 5, 7, 8, 9, 11, 12, 14, 15 and 16), respectively. Gurobi 6.0 is used to solve the ILPs in Section 4.2. The experiments are run on a computer that has an Intel Core(TM) i3-4030U CPU @ 1.90GHz and 4GB memory.

We first provide the comparisons on the maximum amount of data that can be backed up between ILP and heuristic for $|\bar{V}| = 4$ and $|\bar{V}| = 10$, respectively when ε ranges from 1 to 100 time units, as shown in Fig. 4.2. From Fig. 4.2, we can observe that the maximum amount of data that can be backed up achieved by the proposed heuristic is the same as that achieved by ILP. Note that the ILP gives a mathematical formulation for the BCE sub-problem, whereas its optimal solution



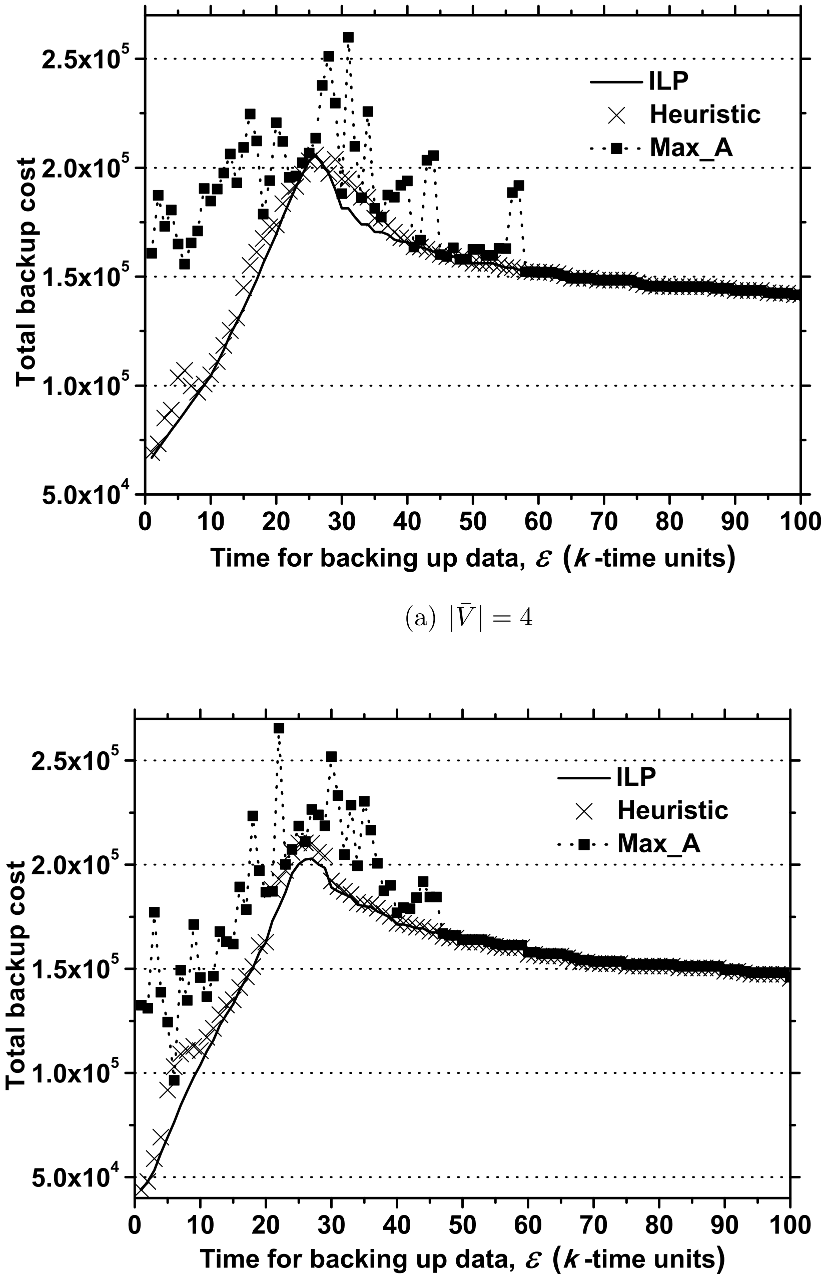
(a) $|\bar{V}| = 4$

(b) $|\bar{V}| = 10$

Figure 4.2: Comparison between ILP and heuristic on the maximum amount of data that can be backed up for different times ε

can indeed be found by the corresponding heuristic. In other words, our heuristic for BCE is an exact algorithm for generating an optimal solution. This is because we assume only a single threatened DCN node. Its maximum amount of protectable data is determined by the available storage capacity of each backup DCN, as well as the available bandwidth on the paths to those backup DCNs. Since our heuristic fully utilizes the available bandwidth on all paths to those backup DCNs, it can exactly achieve an optimal solution as the ILP. Also note that this is only for BCE, and the situation is different for BCM. We can also find that the maximum amount of data that can be backed up increases as the time ε increases and the amount of data that can be backed up reaches to the maximum value 2000 data units for $|\bar{V}| = 4$ when ε is equal to 27 time units and that for $|\bar{V}| = 10$ when ε is equal to 28 time units.

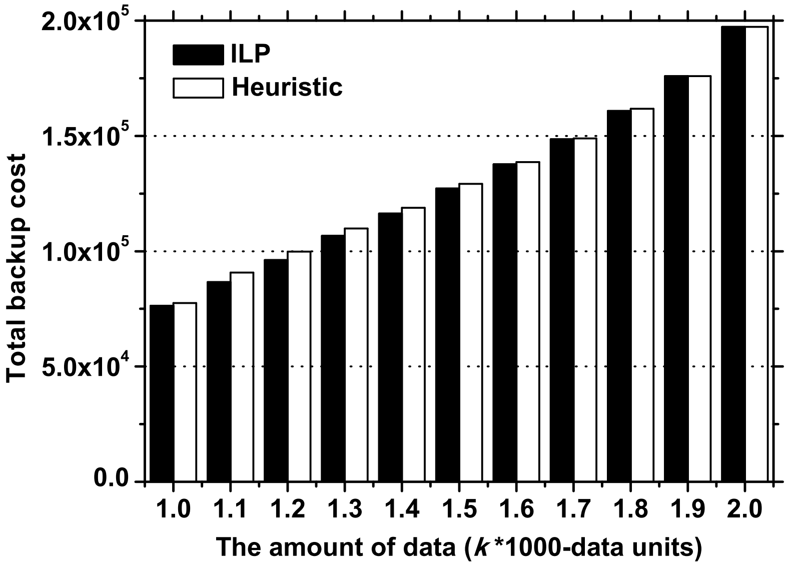
In Fig. 4.3, we then show the total backup cost for each maximum amount of data achieved in Fig. 4.2. For comparison, inspired by references [41] and [43], we also show the results from the backup scheme with the objective of maximizing the amount of data that can be backed up (referred to as Max_A), which is defined here as a benchmark of our proposed scheme in terms of backup cost. From the results in Fig. 4.3, we can observe that Max_A involves the large cost and our proposed scheme is effective in reducing the backup cost. We also use our proposed ILP as a benchmark to evaluate the performance of the proposed heuristic in Fig. 4.3. We can find that the maximum gap between ILP and heuristic is 23.9% for $|\bar{V}| = 4$ when ε is equal to 5 time units and that is 34.9% for $|\bar{V}| = 10$ when ε is equal to 6 time units. The average gap between ILP and heuristic is 5.2% for $|\bar{V}| = 4$ when ε ranges from 1 to 27 time units and that is 8.2% for $|\bar{V}| = 10$ when ε ranges from 1 to 28 time units. The results in Fig. 4.3 also show that after the maximum amount of data that can be backed up reaches to the maximum value 2000 data units, the total backup cost decreases as the time ε increases. This is because more time is available for data backup, and thus less bandwidth is consumed. The above results indicate that the



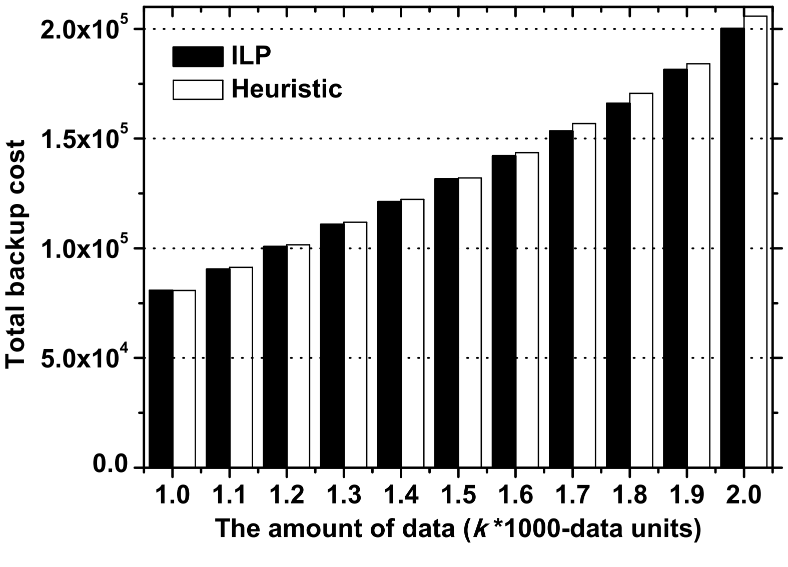
(a) $|\bar{V}| = 4$

(b) $|\bar{V}| = 10$

Figure 4.3: Comparison on the total backup cost of the maximum amount of data that can be backed up based on ILP, heuristic and Max_A for different times ε

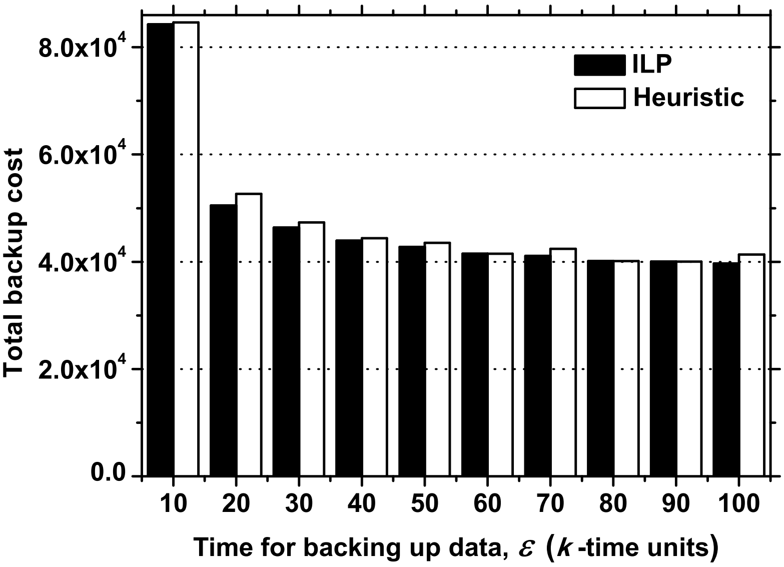


(a) $|\bar{V}| = 4$

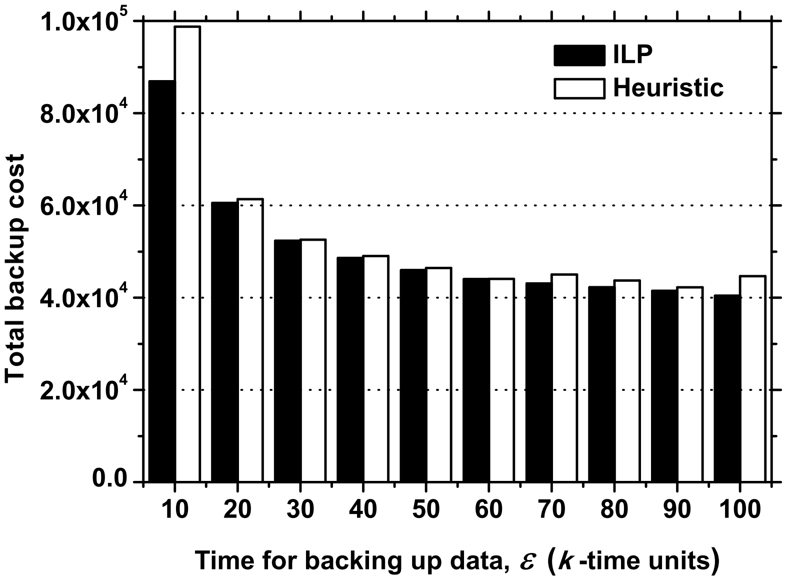


(b) $|\bar{V}| = 10$

Figure 4.4: Total backup cost comparison between ILP and Heuristic for different amounts of data with $\varepsilon = 28$ time units



(a) $|\bar{V}| = 4$



(b) $|\bar{V}| = 10$

Figure 4.5: Total backup cost comparison between ILP and Heuristic for different times ε with $Vl = 700$ data units

Table 4.4: Running time (in second) for solving ILP and executing heuristic with $\varepsilon = 28$ time units

Vl	ILP		Heuristic	
	$\|\bar{V}\| = 4$	$\|\bar{V}\| = 10$	$\|\bar{V}\| = 4$	$\|\bar{V}\| = 10$
1000	3.072	11.587	0.168	0.281
1100	4.066	8.178	0.049	0.984
1200	5.123	7.546	0.056	0.13
1300	5.649	7.391	0.062	0.078
1400	3.041	4.75	0.037	0.21
1500	3.541	6.03	0.903	0.103
1600	2.695	5.822	0.033	0.057
1700	3.485	8.429	0.036	0.08
1800	4.254	9.845	0.034	0.065
1900	3.369	72.342	0.047	0.093
2000	1.717	3.723	0.036	0.063

proposed heuristic is efficient.

To further validate the performance of the proposed heuristic, we also give the following comparisons of the ILP and heuristic. Fig. 4.4 shows total backup costs from ILP and heuristic for $|\bar{V}| = 4$ and $|\bar{V}| = 10$, respectively when Vl ranges from 1000 to 2000 data units at a fixed $\varepsilon = 28$ time units. The results in Fig. 4.4 indicate that total backup cost increases with the increase of Vl. Although the gap of backup cost between ILP and heuristic varies as Vl and $|\bar{V}|$ increase, the gap is always less than 5%. In Fig. 4.5, we show total backup costs from ILP and heuristic for $|\bar{V}| = 4$ and $|\bar{V}| = 10$, respectively when we increase ε from 10 to 100 time units at a fixed Vl=700 data units. The results in Fig. 4.5 indicate that the total backup cost decreases as ε increases. We also find that the gap of backup cost between ILP and heuristic is less than 14%. The above results also indicate that the proposed heuristic is efficient. It is notable that the results from Figs. 4.2, 4.3 and 4.5 show that our proposed scheme can automatically adapt to disasters with different early warning times ε for generating efficient data backup solutions.

Tables 4.4 and 4.5 show running times for solving ILP and executing heuristic

Table 4.5: Running time (in second) for solving ILP and executing heuristic with Vl = 700 data units

ε	ILP		Heuristic	
	$\|\bar{V}\| = 4$	$\|\bar{V}\| = 10$	$\|\bar{V}\| = 4$	$\|\bar{V}\| = 10$
10	5.224	5.969	0.197	0.312
20	1.411	3.683	0.056	1.027
30	3.525	5.138	0.048	0.087
40	1.917	2.635	0.061	0.082
50	2.159	2.867	1.077	0.083
60	1.505	3.069	0.039	0.081
70	1.442	2.502	0.04	0.052
80	2.942	4.414	0.036	0.18
90	1.519	3.088	0.036	0.058
100	1.633	2.458	0.03	0.126

in Figs. 4.4 and 4.5, respectively. We can observe that the time for solving ILP increases with the increase of $|\bar{V}|$. In particular, the time for solving ILP reaches to the maximum value that more than 72 seconds when $Vl = 1900$ and $|\bar{V}| = 10$. However, the time for executing heuristic increases slowly with the increases of $|\bar{V}|$ and Vl and thus the proposed heuristic is more scalable. Since the time for executing heuristic is small for large-scale backup problems, we can achieve a real-time solution based on the proposed heuristic to meet the practical engineering requirement against an ε-time early warning disaster. For example, under the above mentioned hardware settings (i.e., an Intel Core(TM) i3-4030U CPU @ 1.90GHz and 4GB memory), the proposed heuristic can provide backup schemes for all the scenarios in Fig. 4.4 against a disaster with ε=29 time units early warning time.

4.5 Summary

We studied the minimum cost data backup in geo-distributed DCNs against an ε-time early warning disaster under a given set of backup resources. Two sets of algorithms were proposed, each consisting of an optimal ILP and a corresponding heuristic. With the ε early warning time constraint, the first set of algorithms can help DCN operators to evaluate the maximum backup capacity under the limited amount of backup resources, and the second set of algorithms can minimize backup cost by properly selecting a set of backup DCN nodes and corresponding backup routes. By properly exploring the ε early warning time, the proposed scheme can be more flexible and adaptive to disasters as compared with existing periodical backup and real-time replication schemes. Our scheme allows simultaneous data backup from the threatened DCN node to multiple safe DCN nodes in the disaster-disjoint zones. It was shown that the optimal solution changes with different early warning times ε, indicating that the proposed scheme is disaster adaptive under different values of ε.

CHAPTER V

Heterogeneous Data Backup Based on Early Warning of Region Failure

In this chapter, we investigate the heterogeneous data backup based on early warning of region failure. Similar to Chapter IV, we also assume that there is only one data center network (DCN) node (i.e., threatened DCN node) falling within the region that will be affected by a disaster after ε early warning time. We consider urgent backup within the early warning time ε where different types of data at the threatened DCN node may be backed up to the different sets of backup DCN nodes. To this end, by fully utilizing the ε early warning time, we propose two backup schemes which are maximum data backup scheme (MDBS) and fairness data backup scheme (FDBS). The former is to maximize the total amount of data that can be backed up, and the latter is to maximize the same proportion of data backup for each type of data in a fair manner. Corresponding integer linear program (ILP) and heuristic are developed for each scheme. Extensive numerical results show that the proposed schemes can flexibly provide different data backup solutions against the disasters with different early warning times. Note that the allowed storage manner of data in this chapter is similar to that in Chapter IV.

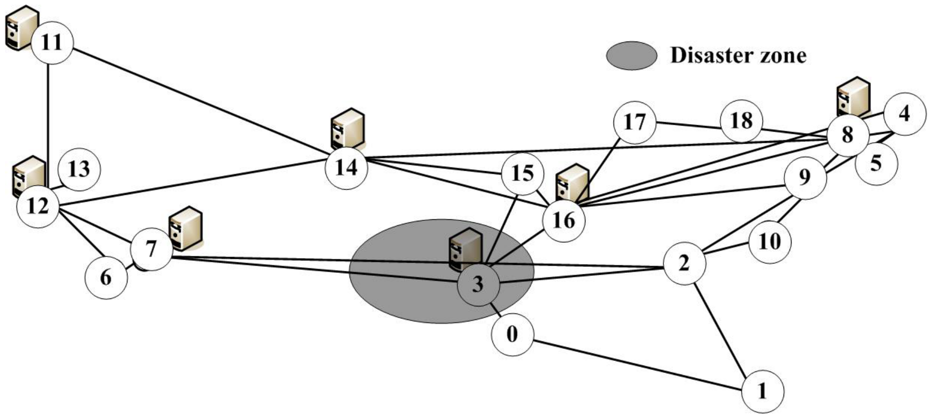


Figure 5.1: Illustration of one threatened DCN node and six candidate backup DCN nodes in geo-distributed DCNs under disaster

5.1 Network Model

We consider the similar network scenario as that in Chapter IV. The network topology is also modeled as a graph $G(V, E)$, where V is the set of all nodes and E is the set of all fiber links. We also assume that there is a single threatened DCN node in the network. The backup DCN nodes are selected from the other safe DCN nodes under disaster. Each candidate backup DCN node has a certain amount of backup storage, whereas the available wavelengths on each network link for backup routing can be measured by the DCN operator at the disaster time. Fig. 5.1 gives an example of the U.S. InternetMCI network [70] with seven geo-distributed DCNs hosted at nodes 3, 7, 8, 11, 12, 14 and 16. Suppose there is a sudden disaster will affect the DCN node 3 area after ε early warning time, as shown by the shaded area in Fig. 5.1 and there are four types of data that should be backed up in the DCN node 3. Then, four different backup requirements are formed in such network, i.e., {data ID, {backup DCN node list}} (e.g., {1, {8, 12}}, {2, {7, 14}}, {3, {11, 16}}, {4, {7, 8}}). Note that similar to Chapter IV, store-and-forward schemes are also not considered in this chapter.

5.2 ILP Formulations

In this section, we provide two ILPs for MDBS and FDBS under the early warning time constraint, respectively.

5.2.1 Notation List

The detailed inputs and variables used in the ILP formulations are listed in Tables 5.1 and 5.2.

Table 5.1: Parameters for Inputs

Notation	Definition
V	Similar to the definition in Section 3.2.2.
E	Similar to the definition in Section 3.2.2.
$\bar{V} \subset V$	Similar to the definition in Section 4.2.1.
ε	Similar to the definition in Section 4.2.1.
$P = \{p \vert p =< Sc_p, De_p, L_p >\}$	Similar to the definition in Section 4.2.1.
Re	Similar to the definition in Section 4.2.1.
S_v	Similar to the definition in Section 4.2.1.
B_e	Similar to the definition in Section 4.2.1.
$A_e^p \in \{0, 1\}$	Similar to the definition in Section 4.2.1.
$D' = \{d \vert d =< C_d, Bun_d, P_d >\}$	The set of different types of data at the threatened DCN node, where C_d is the amount of the type of data d that should be backed up (It is quantified with the number of data units), $Bun_d \subseteq \bar{V}$ is a set of candidate backup DCN nodes for backing up the type of data d and $P_d \subseteq P$ is a set of possible paths for backing up the type of data d.

χ	Predefined constant greater than $max\{B_p^d, Su_v^d \mid \forall d \in D', \forall v \in Bun_d, \forall p \in P_d\}$.

Table 5.2: Variables

Notation	Definition
U_v^d	Binary variable. It takes 1 if the DCN node $v \in Bun_d$ is used for backing up the type of data $d \in D'$ and 0 otherwise.
U_p^d	Binary variable. It takes 1 if path $p \in P_d$ is used for backing up the type of data $d \in D'$ and 0 otherwise.
Su_v^d	Non-negative integer. It is the used storage capacity in DCN node $v \in Bun_d$ for backing up the type of data $d \in D'$.
B_p^d	Non-negative integer. It is the used bandwidth on path $p \in P_d$ for backing up the type of data $d \in D'$.
θ	Non-negative integer ($0 < \theta \leq N, N = 10^n$, integer $n \geq 2$). $\frac{\theta}{N}$ is the same proportion of data backup for each type of data in set D'.

5.2.2 ILP for Maximum Data Backup Scheme

$$Maximize \quad \Big\{ \sum_{d \in D'} \sum_{v \in Bun_d} Su_v^d \Big\}. \tag{5.1}$$

Subject to

$$\sum_{v \in Bun_d} Su_v^d \leq C_d, \forall d \in D'; \tag{5.2}$$

$$\sum_{d \in D'} Su_v^d \leq S_v, \forall v \in \bar{V}; \tag{5.3}$$

$$\sum_{d \in D'} \sum_{p \in P_d} A_e^p B_p^d \leq B_e, \forall e \in E; \tag{5.4}$$

$$\sum_{v \in Bun_d} U_v^d \geq 1, \forall d \in D'; \tag{5.5}$$

$$U_p^d \leq \frac{U_{De_p}^d + 1}{2}, \forall d \in D', \forall De_p \in Bun_d, \forall p \in P_d; \tag{5.6}$$

$$\sum_{p \in P_d, De_p = v} U_p^d \geq U_v^d, \forall d \in D', \forall v \in Bun_d; \tag{5.7}$$

$$U_p^d \leq B_p^d, \forall d \in D', \forall p \in P_d; \tag{5.8}$$

$$U_p^d \geq B_p^d / \chi, \forall d \in D', \forall p \in P_d; \tag{5.9}$$

$$U_v^d \leq Su_v^d, \forall d \in D', \forall v \in Bun_d; \tag{5.10}$$

$$U_v^d \geq Su_v^d/\chi, \forall d \in D', \forall v \in Bun_d; \tag{5.11}$$

$$\frac{Su_v^d}{\sum\limits_{p \in P_d, De_p = v} B_p^d} \leq \varepsilon \cdot Re, \forall d \in D', \forall v \in Bun_d. \tag{5.12}$$

Objective (5.1) maximizes the total amount of data that can be backed up at the threatened DCN node. Constraint (5.2) limits the total amount of each type of data that is backed up to its maximum possible amount. Constraint (5.3) ensures that the used storage capacity for backing up data in each backup DCN node $v \in \bar{V}$ does not exceed the available storage capacity of such DCN node. Constraint (5.4) ensures that the used bandwidth for backing up data on a link does not exceed the available capacity on this link. Constraint (5.5) guarantees that each type of data is backed up to at least one backup DCN node. Constraint (5.6) implies that if a path $p \in P_d$ is selected for backing up the type of data $d \in D'$, then the destination node of this path $De_p \in Bun_d$ must be selected as the backup DCN node for storing such type of data. Constraint (5.7) implies that if a DCN node is selected as the backup DCN node for backing up the type of data $d \in D'$, then at least one path must be selected as the transmission path for backing up this type of data in such DCN node. Constraints (5.8) and (5.9) define U_p^d whereas constraints (5.10) and (5.11) define U_v^d. Constraint (5.12) ensures that the time for backing up each type of data does not exceed the given backup time ε.

5.2.3 ILP for Fairness Data Backup Scheme

The objective of the ILP for FDBS is to maximize θ as formulated in (5.13), such that the same proportion of data backup for each type of data $d \in D'$ (i.e., $\frac{\theta}{N}$) is maximized.

$$Maximize \quad \{\theta\}. \tag{5.13}$$

The constraints of the ILP include constraints (5.3)-(5.12) in the ILP for MDBS and the following new constraints (5.14) and (5.15) which determine the amount of each type of data that should be backed up.

$$\sum_{v \in Bun_d} Su_v^d \geq \frac{\theta}{N} \cdot C_d - 1, \forall d \in D'; \tag{5.14}$$

$$\sum_{v \in Bun_d} Su_v^d \leq \frac{\theta}{N} \cdot C_d, \forall d \in D'. \tag{5.15}$$

Since the amount of each type of data that can be backed up is set as an integer, there may be a gap between $\sum\limits_{v \in Bun_d} Su_v^d$ and $\frac{\theta}{N} \cdot C_d$ for each type of data $d \in D'$. However, the gap will not be greater than 1. Note that the assumptions of wavelength converters at intermediate nodes (if necessary) in the network and static network status within the early warning time are the same as those assumed in Chapter IV.

5.3 Heuristics

Since solving ILP for large-scale problems (e.g. more types of data that should be backed up) induces the high time complexity, it is generally difficult to obtain an optimal solution based on ILP for data backup within a limited time. To make our schemes more scalable, in this section we propose time-efficient heuristics for MDBS (i.e., Algorithm 1) and FDBS (i.e., Algorithm 2) to get the real-time solutions, respectively. In these algorithms, $\bar{V}$, Re, D', P, S_v, B_e, ε, n and N are defined in

```
Algorithm 1 Maximum Data Backup Scheme (MDBS):
Input:
    $G(V,E)$, $\bar{V} \subset V$, $Re$, $D'$, $P$, $S_v$ for $\forall v \in \bar{V}$, $B_e$ for $\forall e \in E$, and the time for backing up data $\varepsilon$.
Output:
    The backup scheme, i.e., the sets of backup DCN nodes ($V_b^d$) and backup transmission paths ($T_p^d$) for each type of data $d \in D'$, and the amount of each type of data $d \in D'$ that can be backed up within the time $\varepsilon$, $M_d$.
1: $V_b^d = \phi$, $T_p^d = \phi$ , $M_d = 0$ for $\forall d \in D'$, $P' = P$, $Flag = 1$;
2: Call Procedure 1 Data Backup A;
3: Call Procedure 2 Data Backup B.
```

Section 5.2.1, and the bandwidth on transmission path is assigned with the integer. We also use $|\Phi|$ to denote the number of elements in a given set Φ.

5.3.1 Heuristic for Maximum Data Backup Scheme

The proposed heuristic for MDBS is illustrated in Algorithm 1. To achieve the maximum amount of data that can be backed up, the iterative method is adopted. In each iteration, for each path p in set P we calculate the amount of data that can be backed up through path p within the given early warning time ε, where the amount of data that can be backed up through path p depends on the type of data with the maximum amount at the threatened DCN node that can be backed up to the destination node of path p, the available storage capacity in the destination node of path p and the available bandwidth on path p. Then we can select a path that can be used to back up the maximum amount of data from set P to execute backup. In Algorithm 1, we first set $V_b^d = \varnothing$, $T_p^d = \varnothing$, $M_d = 0$, $P' = P$ and $Flag = 1$ in line 1. Then the Procedure 1 (i.e., Data Backup A) is executed by taking the inputs of Algorithm 1 as its inputs. After executing Procedure 1, we call Procedure 2 (i.e., Data Backup B) by taking the updated inputs of Algorithm 1 by Procedure 1 as its inputs.

Data Backup A: In this procedure, we back up the amount of data $\varepsilon \cdot Re \cdot B_t$,

Procedure 1 Data Backup A:

```
1: while ($Flag = 1$) do
2:     Select a path $p$ with the nonzero maximum available bandwidth to back up data based on (5.16) from the set $P'$;
3:     if ($p$ is found) then
4:         Determine a bandwidth $B_p^{d_{max}^p}$ on path $p$ for backing up the type of data $d_{max}^p$ based on (5.16);
5:         Set $C_{d_{max}^p} = C_{d_{max}^p} - B_p^{d_{max}^p} \cdot \varepsilon \cdot Re$, $S_{De_p} = S_{De_p} - B_p^{d_{max}^p} \cdot \varepsilon \cdot Re$;
6:         Set $B_e = B_e - B_p^{d_{max}^p}$ for $\forall e \in p$;
7:         Set $V_b^{d_{max}^p} = V_b^{d_{max}^p} \bigcup De_p$, $T_p^{d_{max}^p} = T_p^{d_{max}^p} \bigcup p$;
8:         Set $M_{d_{max}^p} = M_{d_{max}^p} + B_p^{d_{max}^p} \cdot \varepsilon \cdot Re$;
9:     else
10:        Set Flag=0;
11:    end if
12: end while
```

where B_t denotes the total number of wavelengths that are selected for backing up data. For a path p that is selected to back up data, we use d_{max}^p to denote the type of data with the maximum amount $C_{d_{max}^p}$ in set D' that can be backed up to the backup DCN node De_p. In line 2, we select a path p with the nonzero maximum available bandwidth to back up data from set P' (i.e., which can be used to back up the maximum amount of data). Here, the available bandwidth on path p is $Min_{e \in p}\{B_e\}$ and the maximum available bandwidth on path p for backing up data is determined as

$$Min\Big(\Big\lfloor \frac{S_{De_p}}{\varepsilon \cdot Re} \Big\rfloor, \Big\lfloor \frac{C_{d_{max}^p}}{\varepsilon \cdot Re} \Big\rfloor, Min_{e \in p}\{B_e\}\Big). \tag{5.16}$$

The above expression (5.16) ensures that the assigned bandwidth on path p for backing up data d_{max}^p satisfies the constraints of the available capacity of DCN node De_p, the amount of data $C_{d_{max}^p}$ and the available bandwidth on path p. If we can find an available path p, in line 4 the bandwidth $B_p^{d_{max}^p}$ on path p for backing up data d_{max}^p is obtained based on (5.16). We update the values of $C_{d_{max}^p}$, S_{De_p} and B_e for each $e \in p$ in lines 5-6, respectively and then we add the node De_p and path p into the sets

Procedure 2 Data Backup B:

```
1: while ($Flag = 0$) do
2:     Select a path $p$ from the set $P'$ with nonzero available bandwidth $Min_{e \in p}\{B_e\}$
       which can be used to back up the nonzero maximum amount of data based on
       (5.17);
3:     if ($p$ is found) then
4:         Determine the amount of the type of data $d^p_{max}$ that can be backed up through
           path $p$ based on (5.17);
5:         Determine a bandwidth $B_p^{d^p_{max}} = 1$ on path $p$ for backing up the type of data
           $d^p_{max}$ ;
6:         Set $C_{d^p_{max}} = C_{d^p_{max}} - Min(C_{d^p_{max}}, S_{De_p})$, $S_{De_p} = S_{De_p} - Min(C_{d^p_{max}}, S_{De_p})$;
7:         Set $B_e = B_e - B_p^{d^p_{max}}$ for $\forall e \in p$;
8:         Set $V_b^{d^p_{max}} = V_b^{d^p_{max}} \bigcup De_p$, $T_p^{d^p_{max}} = T_p^{d^p_{max}} \bigcup p$;
9:         Set $M_{d^p_{max}} = M_{d^p_{max}} + Min(C_{d^p_{max}}, S_{De_p})$;
10:    else
11:        Set Flag=1;
12:    end if
13: end while
```

$V_b^{d^p_{max}}$ and $T_p^{d^p_{max}}$ in line 7, respectively. The total amount of data d^p_{max} (i.e., $M_{d^p_{max}}$) that is backed up to the backup DCN nodes is obtained in line 8. If we can not find an available path p, the procedure exits.

Data Backup B: In this procedure, we back up the remaining amount of each type of data in set D'. In line 2, we select a path p with nonzero available bandwidth from set P' which can be used to back up the nonzero maximum amount of data.

$$Min\Big(C_{d^p_{max}}, S_{De_p}\Big). \tag{5.17}$$

If we can find an available path p, the maximum amount of data that can be backed up through path p is determined by (5.17) in line 4 and the bandwidth on path p for backing up the type of data d^p_{max} is set as 1 in line 5. Then, in lines 6-9 we can execute the similar operations as those in lines 5-8 of Procedure 1. If we can not find an available path p, the procedure exits.

Complexity of the heuristic: In Algorithm 1, the complexity of the operation in line 1 is $O(Max(|D'| + |P|))$. Then we analyze the complexity of Procedure 1. In Procedure 1, the while-loop from lines 1-12 is executed at most $|P|$ times for backing up data, i.e., we need to use each path in set P' to back up data in the worst case scenario. In order to select a path p from set P' in line 2, we need to traverse all available paths for backing up data in set P', and for each traversed path p, we need to traverse all types of data in set D' that can be backed up to the backup DCN node De_p. Thus, the complexity of line 2 is no more than $O(|P| \times |E| \times |D'|)$. The time complexity for update from lines 5-8 is $O(|E|)$. Thus, the time complexity of Procedure 1 is no more than $O(|P|^2 \times |E| \times |D'|)$. Besides, since Procedure 2 has the same complexity of Procedure 1, the overall time complexity of the Algorithm 1 is $O(|P|^2 \times |E| \times |D'|)$.

5.3.2 Heuristic for Fairness Data Backup Scheme

The proposed heuristic for FDBS is illustrated in Algorithm 2. To achieve the data backup for each type of data $d \in D'$ in a fair manner, we maximize the same proportion of data backup for each type of data in set D' within the early warning time ε (X) by using the idea of binary search. For each type of data $d \in D'$, the backup operation is implemented by the Procedures 1 and 2. When we back up the type of data $d \in D'$, the set P' and the data $d^p_{max}, p \in P'$ in those procedures are set as P_d and d, respectively. In Algorithm 2, the initialization is first shown in lines 1-4, where we use C'_d, S'_v and B'_e to save the initial values of the amount of the type of data $d \in D'$, the available storage capacity of the backup DCN node $v \in \bar{V}$ and the available bandwidth on link $e \in E$, respectively. The while-loop from lines 5-28 executes the binary search to find the maximum proportion X. In each loop, we first execute the initialization operations from lines 6-11 where the amount of each type of data $d \in D'$ that should be backed up is determined by the proportion $\frac{\theta}{N}$. Then, according to the determined amount of each type of data $d \in D'$, we sort different

Algorithm 2 Fairness Data Backup Scheme (FDBS):

```
Input:
    $G(V, E)$, $\bar{V} \subset V$, $Re$, $D'$, $P$, $S_v$ for $\forall v \in \bar{V}$, $B_e$ for $\forall e \in E$, the integer $N$ ($n \geq 2$) and the time for backing up data $\varepsilon$.
Output:
    The backup scheme, i.e., the sets of backup DCN nodes ($V_b^d$) and backup transmission paths ($T_p^d$) for each type of data $d \in D'$, the amount of each type of data $d \in D'$ that can be backed up within the time $\varepsilon$, $M_d$, and the maximum same proportion of data backup for each type of data in set $D'$ within the time $\varepsilon$, $X$.
Set $low = 0$, $high = N$;
Set $C'_d = C_d$ for $\forall d \in D'$;
Set $S'_v = S_v$ for $\forall v \in \bar{V}$;
Set $B'_e = B_e$ for $\forall e \in E$;
while ($low <= high$) do
    Set $S_v = S'_v$ for $\forall v \in \bar{V}$;
    Set $B_e = B'_e$ for $\forall e \in E$;
    Set $C_d = C'_d$, $V_b^d = \phi$ and $T_p^d = \phi$ for $\forall d \in D'$;
    Set $\theta = \lfloor (low + high)/2 \rfloor$, $Flag = 1$;
    Set $C_d = \lfloor \frac{C_d * \theta}{N} \rfloor$ for $\forall d \in D'$;
    Set $Total = \sum_{d \in D'} C_d$;
    Sort different types of data in set $D'$ in a descending order according to the amount of each type of data (i.e., $C_d, \forall d \in D'$);
    for ($\forall d \in D'$) do
        Set $P' = P_d$;
        Call Procedure 1 Data Backup A;
    end for
    Sort different types of data in set $D'$ in a descending order according to the remaining amount of each type of data (i.e., $C_d, \forall d \in D'$);
    for ($\forall d \in D'$) do
        Set $P' = P_d$;
        Call Procedure 2 Data Backup B;
    end for
    if ($Total = \sum_{d \in D'} M_d$) then
        Set $low = \theta + 1$;
        Set $\theta' = \theta$;
    else
        Set $high = \theta - 1$;
    end if
end while
Set $X$=$Min_{d \in D'}(\lfloor \frac{C'_d * \theta'}{N} \rfloor / C'_d)$.
```

Procedure 3 Data Backup C:

```
for each path $p \in P'$ do
    Determine a bandwidth $B_p^{d_{max}^p}$ on path $p$ for backing up the type of data $d_{max}^p$ based on (5.18);
    if $(B_p^{d_{max}^p} > 0)$ then
        Set $C' = Min(C_{d_{max}^p}, S_{De_p}, B_p^{d_{max}^p} \cdot \varepsilon \cdot Re)$;
        Set $C_{d_{max}^p} = C_{d_{max}^p} - C'$, $S_{De_p} = S_{De_p} - C'$;
        Set $B_e = B_e - B_p^{d_{max}^p}$ for $\forall e \in p$;
        Set $V_b^{d_{max}^p} = V_b^{d_{max}^p} \bigcup De_p$, $T_p^{d_{max}^p} = T_p^{d_{max}^p} \bigcup p$;
        Set $M_{d_{max}^p} = M_{d_{max}^p} + C'$;
    end if
end for
```

types of data in set D' in a descending order in line 12. Following, for each type of data in set D', the backup operation is executed by calling Procedure 1 in lines 13-16. From lines 17-21, the similar operations are executed as those in lines 12-16 for the remaining amount of each type of data $d \in D'$, where Procedure 2 is called. After that if the determined amount of each type of data $d \in D'$ in line 10 can be backed up within the time ε, we change the search scope from $\theta + 1$ to *high* in line 23. Otherwise, we change the search scope from *low* to $\theta - 1$ in line 26. At last, the maximum proportion X can be obtained in line 29.

Complexity of the heuristic: In Algorithm 2, the complexity of the initialization operations from lines 1-4 is $O(Max(|D'|, |\bar{V}|, |E|))$. The while-loop from lines 5-28 is executed $log_2 N (N = 10^n)$ times. In the loop, the complexity of the operations from lines 6-11 is $O(Max(|D'|, |\bar{V}|, |E|))$ and the complexity for sorting data is at most $O(|D'|^2)$. Then, since both of the complexities of Procedures 1 and 2 for solving FDBS are $O(|P|^2 \times |E|)$, the overall complexity of Algorithm 2 is $O(log_2 N(|D'|^2 + Max(|D'|, |\bar{V}|, |E|) + |D'| \times |P|^2 \times |E|))$.

Table 5.3: Backup requirements

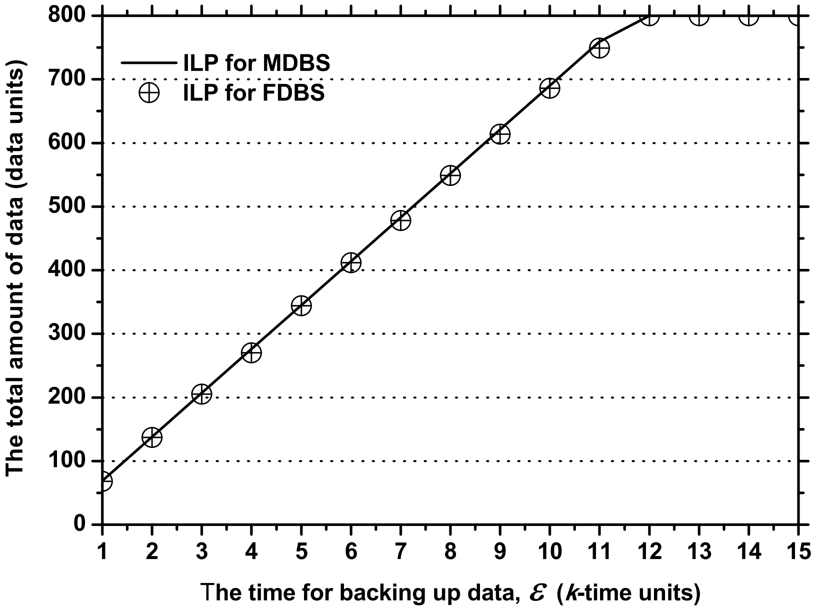
$\lvert D'\rvert = 4$		$\lvert D'\rvert = 8$	
Data ID	Backup DCNs	Data ID	Backup DCNs
1	8, 12	1	8, 12
2	7, 14	2	7, 14
3	11, 16	3	11, 16
4	7, 12	4	7, 12
		5	12, 16
		6	7, 14
		7	14, 16
		8	8, 11

5.4 Numerical Results

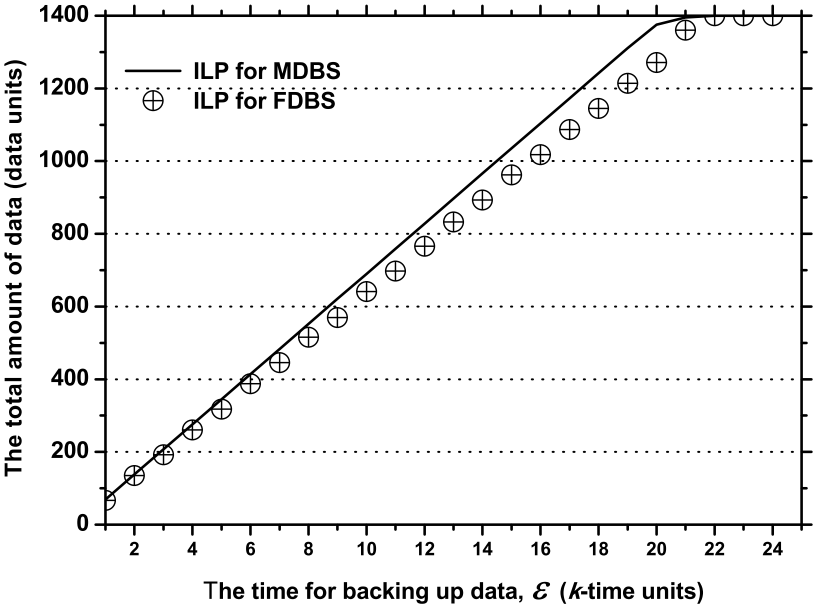
In this section, we simulate MDBS and FDBS on U.S. InternetMCI network to implement urgent backup. As shown in Fig. 5.1, the network includes 19 nodes and 33 links, and there are seven DCNs hosted at the network. We assume that each link in the network has a number of available wavelength channels which is determined by randomly generating an integer between 10 and 30. We also assume that an upcoming disaster will affect the DCN node 3 area and different types of data at this node should be backed up within the ε early warning time, where the amount of each type of data is set as a random integer between 100 and 300 data units.

In our experiments, we set $\chi = 1000$, $Re = 1$ and $n = 2$. We consider two scenarios (i.e., $|D'| = 4$ and $|D'| = 8$) for the backup problem, respectively. Here, the total available storage capacity in all backup DCN nodes is set as 1400 data units for $|D'| = 4$ and 2000 data units for $|D'| = 8$. The backup requirements from different types of data for $|D'| = 4$ and $|D'| = 8$ are shown in Table 5.3. Gurobi 6.0 is used to solve the ILPs in Section 5.2. The experiments are run on a computer that has an Intel Core(TM) i3-4030U CPU @ 1.90GHz and 4GB memory.

In Fig. 5.2, we first show the results on the maximum amount of data that can be backed up obtained by MDBS and FDBS based on ILPs for $|D'| = 4$ and

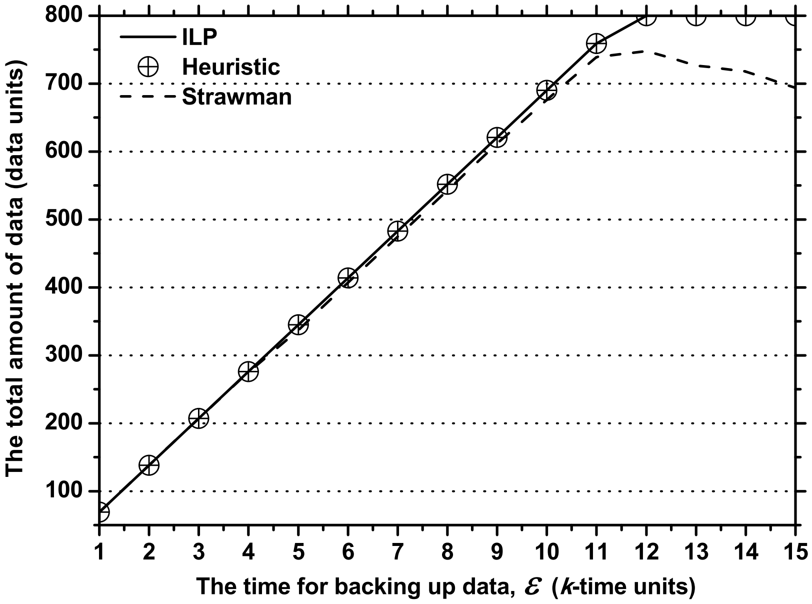


(a) $|D'| = 4$

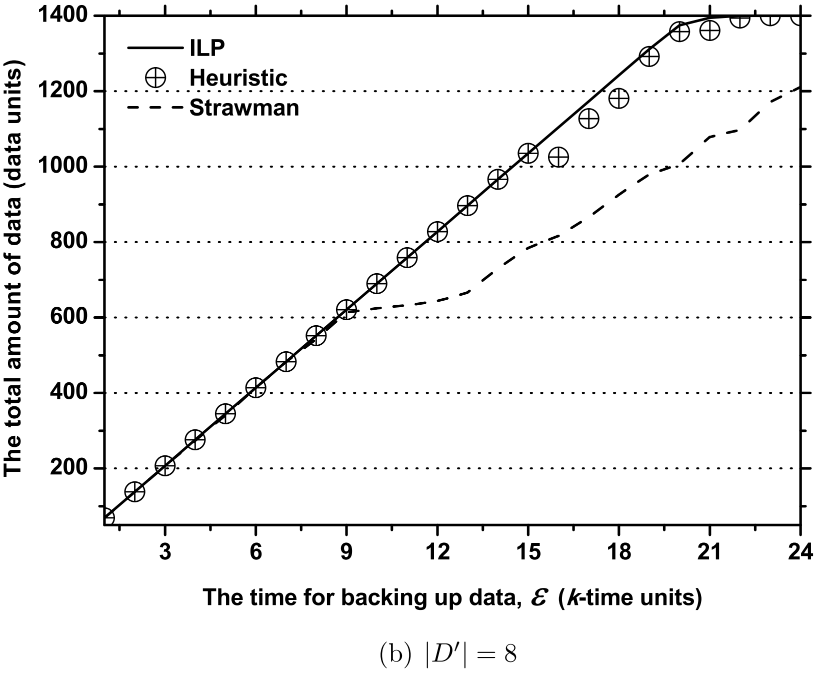


(b) $|D'| = 8$

Figure 5.2: Comparison between MDBS and FDBS on the maximum amount of data that can be backed up based on ILPs for different times ε

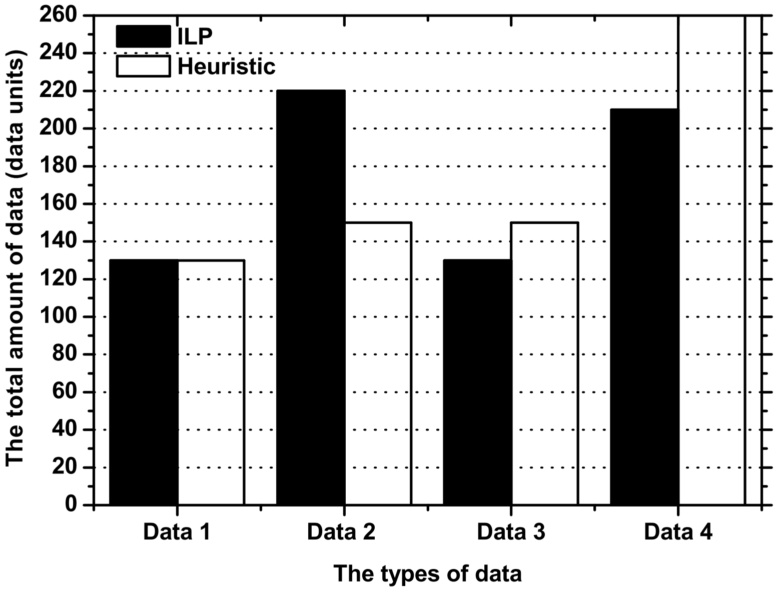


(a) $|D'| = 4$

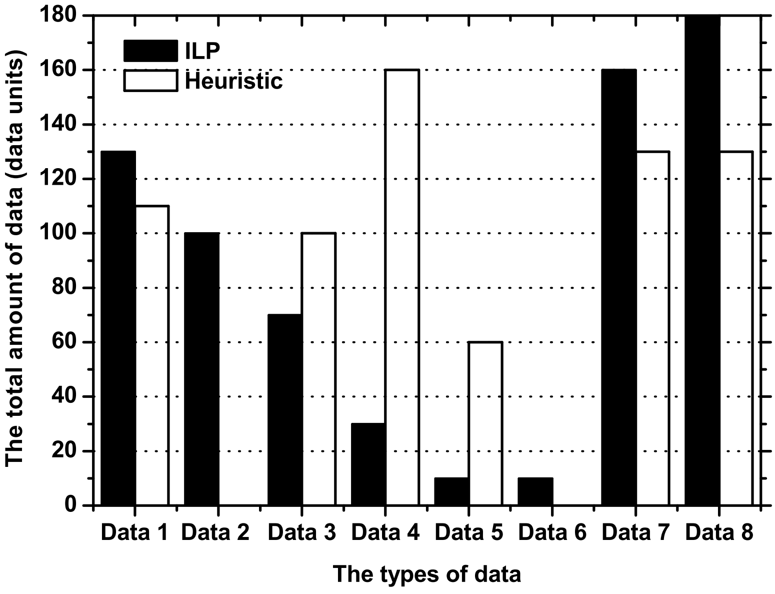


(b) $|D'| = 8$

Figure 5.3: Comparison on the total amount of data that can be backed up from MDBS based on ILP, heuristic and strawman for different times ε



(a) $|D'| = 4$



(b) $|D'| = 8$

Figure 5.4: Comparison between ILP and heuristic on the amount of each type of data that can be backed up from MDBS with $\varepsilon = 10$ time units

$|D'| = 8$, respectively under different times ε. The results in Fig. 5.2 indicate that the maximum amount of data that can be backed up increases with the increase of the time ε. We can also observe that the maximum amount of data that can be backed up obtained by FDBS is less than (or equal to) that obtained by MDBS in both cases, and the gap between MDBS and FDBS increases with the increase of $|D'|$. This is because that the objective of MDBS is to maximize the total amount of data that can be backed up by fully utilizing the available resources and then the upper bound of the amount of data that can be backed up within the given early warning time can be achieved, but FDBS needs to ensure the same proportion of data backup for each type of data which may not fully utilize the available resources and thus the upper bound of the amount of data that can be backed up cannot be always reached for any case based on FDBS.

To validate the effectiveness of our proposed heuristics for MDBS and FDBS, we also provide simple algorithms to solve MDBS and FDBS (referred to as strawman algorithms), respectively. Strawman algorithms are obtained by only calling Procedure 3 in Algorithms 1 and 2, and removing the sort operation in Algorithm 2 as well. In such Procedure 3, we successively select an available path from set P to back up data, and the bandwidth on path p for backing up data is determined by (5.18). The other operations are similar to those in Procedures 1 and 2.

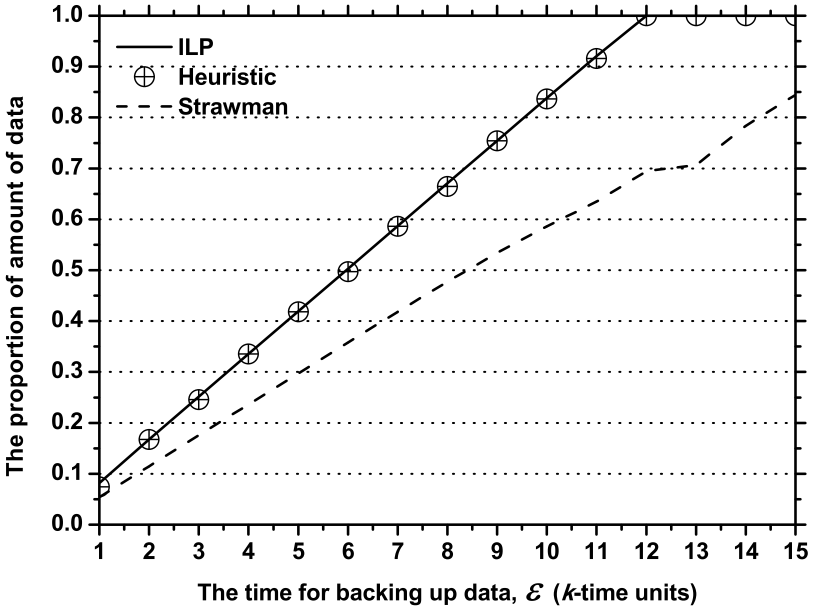
$$Min\Big(\Big\lceil\frac{S_{De_p}}{\varepsilon \cdot Re}\Big\rceil, \Big\lceil\frac{C_{d^p_{max}}}{\varepsilon \cdot Re}\Big\rceil, Min_{e \in p}\{B_e\}\Big). \tag{5.18}$$

We then provide the results obtained by MDBS based on ILP, heuristic and s-trawman algorithm for $|D'| = 4$ and $|D'| = 8$, respectively under different times ε, as shown in Fig. 5.3, which use to validate the effectiveness of the proposed heuristic for MDBS. From Fig. 5.3(a), we can find that the total amount of data that can be backed up obtained by heuristic within the given early warning time is the same as that obtained by ILP when $|D'| = 4$. From Fig. 5.3(b), we can observe that there is a gap between ILP and heuristic when $|D'| = 8$, but the maximum gap is less than 7.2% which indicates that the gap varies within a moderate scale. The results in Figs. 5.3(a) and (b) also indicate that the heuristic outperforms the strawman algorithm. Thus, the proposed heuristic for MDBS is efficient.

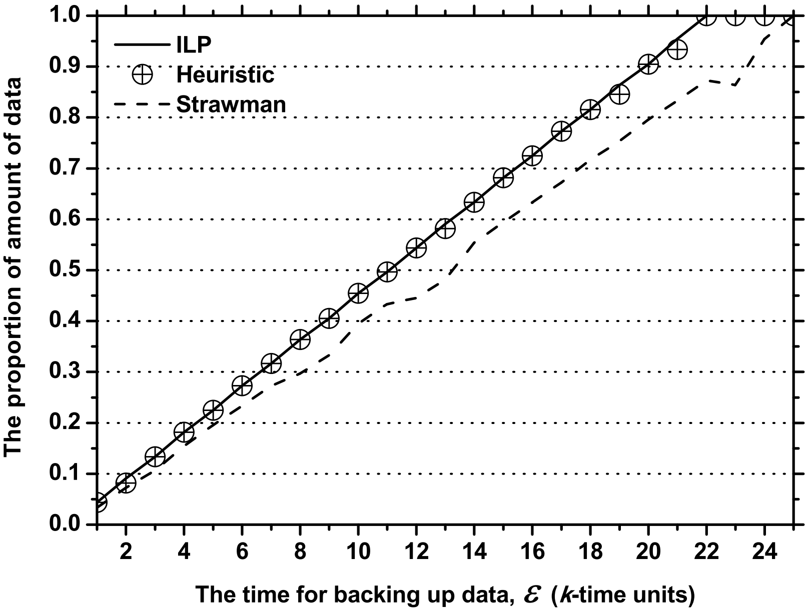
Fig. 5.4 shows the amount of each type of data that can be backed up obtained by MDBS based on ILP and heuristic for two cases (i.e., $|D'| = 4$ and $|D'| = 8$) when $\varepsilon = 10$ time units, respectively. From the results in Fig. 5.4, we can find that although the total amount of data that can be backed up obtained by heuristic is the same as that obtained by ILP for the both cases when $\varepsilon = 10$ time units, the amount of each type of data that can be backed up is different between heuristic and ILP. The results also indicate that MDBS leads to the differential backup for different types of data.

To validate the effectiveness of the proposed heuristic for FDBS, in Fig. 5.5 we also present the results on the same proportion of data backup for each type of data achieved by FDBS based on ILP, heuristic and strawman algorithm for $|D'| = 4$ and $|D'| = 8$, respectively under different times ε. From the results in Fig. 5.5, we can observe that the same proportion of data backup for each type of data obtained by heuristic is similar to that obtained by ILP when $|D'| = 4$, and although the gap of the proportion between ILP and heuristic increases with the increase of $|D'|$, this gap is less than 10% when $|D'| = 8$. We can also find that the heuristic is much better than the strawman algorithm. Thus, the proposed heuristic for FDBS is also efficient. The results in Figs. 5.3 and 5.5 also indicate that the proposed schemes can automatically adapt to different early warning times.

Tables 5.4 and 5.5 show running times for solving ILP and executing heuristic in Figs. 5.3 and 5.5, respectively. From those tables, we can find that the time for



(a) $|D'| = 4$



(b) $|D'| = 8$

Figure 5.5: Comparison on the same proportion of data backup for each type of data from FDBS based on ILP, heuristic and strawman for different times ε

Table 5.4: Running time (in second) for solving ILP and executing heuristic with $|D'| = 4$

ε	ILP		Heuristic	
	MDBS	FDBS	MDBS	FDBS
1	1.476	3.947	0.241	0.326
3	1.231	2.397	0.085	0.15
6	0.708	4.505	0.07	0.056
9	0.6429	5.203	0.132	0.08
12	0.63	1.58	0.116	0.061
15	0.563	1.64	0.055	0.049

Table 5.5: Running time (in second) for solving ILP and executing heuristic with $|D'| = 8$

ε	ILP		Heuristic	
	MDBS	FDBS	MDBS	FDBS
1	5.602	7.871	0.268	0.513
3	6.174	8.285	0.051	0.208
6	3.342	16.454	0.06	0.089
9	3.418	143.646	0.077	0.092
12	5.68	143.368	0.053	0.062
15	3.942	104.311	0.042	0.067
18	2.684	120.072	0.033	0.083
21	> 1000	5.687	0.056	0.059
24	3.79	4.828	0.052	0.059

solving ILP is always larger than that for executing heuristic. We can also observe that the time for solving ILP increases with the increase of $|D'|$. In particular, the time for solving ILP dramatically increases for some given ε when $|D'| = 8$. For example, the time is larger than 100 seconds for FDBS when ε ranges from 9 to 18 time units and that is larger than 1000 seconds for MDBS when $\varepsilon = 21$ time units. However, the time for executing heuristic varies in a small scale. The above results indicate that the proposed heuristics are time-efficient and more scalable, which can provide the real-time solutions against the disaster with a small early warning time.

5.5 Summary

We studied the urgent heterogeneous data backup across geo-distributed DCNs against a disaster with the early warning time under a given set of backup resources. To carry out urgent backup, two backup schemes (i.e., MDBS and FDBS) were proposed, each solving by an optimal ILP and a corresponding heuristic. For different types of data that should be backed up, MDBS can obtain the maximum amount of data that can be backed up, and FDBS can maximize the same proportion of data backup for each type of data to achieve the fair backup for each type of data. Numerical results showed that the proposed schemes can meet the different backup requirements from different types of data and adapt to the disasters with different early warning times.

CHAPTER VI

Conclusion

This chapter summarizes the thesis and points out the interesting future research topics.

6.1 Summary of the Thesis

In this thesis, we focused on the data center network (DCN) placement and data backup against region failures due to disasters. We first investigated the region failure-aware DCN and content placement, and then we explored the homogeneous data backup based on early warning of region failure. Finally, we studied the heterogeneous data backup based on early warning of region failure.

For region failure-aware DCN and content placement, we investigated in Chapter III DCN and content placement with the consideration of non-uniform distribution of region failure in large-scale area. We first evaluated the network vulnerability by integrating the probabilistic region failure model with the grid partition scheme. Based on the network vulnerability information, we then identified the optimal DCN and content placement in the network such that DCN failure probability is minimized, where the solutions for DCN and content placement can be obtained based on ILP and time-efficient heuristic. The results in this chapter showed that our proposed network vulnerability assessment scheme is efficient and also provided the solutions for DCN and content placement under non-uniform spatial and intensity distribution of a potential disaster.

For homogeneous data backup based on early warning of region failure, we proposed in Chapter IV the cost-efficient urgent backup scheme for homogeneous data backup by fully utilizing the ε early warning time. We also provided the scheme to obtain the maximum backup capacity within the ε early warning time. The corresponding ILP models and a heuristic are developed to get the backup solutions. The results in this chapter showed that our proposed schemes are disaster adaptive to different early warning times ε.

In Chapter V, we addressed the heterogeneous data backup based on early warning of region failure. We developed two backup schemes to carry out urgent backup within the ε early warning time. For each scheme, both ILP and heuristic are proposed to generate backup solutions. The results in this chapter also showed that the proposed schemes can flexibly adapt to different early warning times ε.

6.2 Future Work

The interesting future research topics are summarized as follows.

- In this thesis, for DCN and content placement problem, each DCN and each type of content are treated equally. Thus, one interesting future research topic is DCN and content placement with the consideration of different priorities and constraints as well as dynamic network traffic.

- In this thesis, we divide the backup problem into two sub-problems (i.e., Backup Capacity Evaluation (BCE) and Backup Cost Minimization (BCM)) in Chapter IV. Then another attractive future research topic is how to jointly design an optimization problem to maximize the backup capacity while keeping the backup cost minimized.

- It is notable that in heterogeneous data backup, different types of data are treated equally. Thus, it will be interesting topic to develop urgent backup scheme with the consideration of the different priorities for different types of data.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] K. Chen, C. Guo, H. Wu, J. Yuan, Z. Feng, Y. Chen, S. Lu, and W. Wu, “Dac: generic and automatic address configuration for data center networks,” *IEEE/ACM Trans. Netw.*, vol. 20, no. 1, pp. 84–99, Feb. 2012.

[2] D. Abts and B. Felderman, “A guided tour of data-center networking,” *Communications of the ACM*, vol. 55, no. 6, pp. 44–51, Jun. 2012.

[3] M. Bari, R. Boutaba, R. Esteves, L. Granville, M. Podlesny, M. Rabbani, Q. Zhang, and M. Zhani, “Data center network virtualization: a survey,” *IEEE Commun. Surveys Tuts.*, vol. 15, no. 2, pp. 909–928, May 2013.

[4] “Google data centers,” https://www.google.com/about/datacenters.

[5] “Aws global infrastructure,” https://aws.amazon.com/about-aws/global-infrastructure.

[6] “Microsoft global datacenters,” https://www.microsoft.com/en-us/cloud-platform/global-datacenters.

[7] T. Hoff, “Google architecture,” http://highscalability.com/google-architecture, Nov. 2008.

[8] I. Narayanan, A. Kansal, A. Sivasubramaniam, B. Urgaonkar, and S. Govindan, “Towards a leaner geo-distributed cloud infrastructure,” in *2014 6th USENIX Workshop on Hot Topics in Cloud Computing*, 2006, pp. 1–7.

[9] Y. Coady, O. Hohlfeld, J. Kempf, R. McGeer, and S. Schmid, “Distributed cloud computing: applications, status quo, and challenges,” *ACM SIGCOMM Computer Communication Review*, vol. 45, no. 2, pp. 38–43, Apr. 2015.

[10] J. Abley, A. Canada, and K. Lindqvist, *RFC 47867-Operation of Anycast Services*, Dec. 2006.

[11] K. Tanaka, Y. Yamazaki, T. Okazawa, T. Suzuki, T. Kishimoto, and K. Iwata, “Experiment on seismic disaster characteristics of underground cable,” in *The 14th World Conference on Earthquake Engineering*, 2008.

[12] A. Kwasinski, W. W. Weaver, P. L. Chapman, and P. T. Krein, “Telecommunications power plant damage assessment for hurricane katrina-site survey and follow-up results,” *IEEE Systems Journal*, vol. 3, no. 3, pp. 277–287, Sep. 2009.

[13] Y. Ran, "Considerations and suggestions on improvement of communication network disaster countermeasures after the wenchuan earthquake," *IEEE Commun. Mag.*, vol. 49, no. 1, pp. 44–47, Jan. 2011.

[14] K. Morrison, "Rapidly recovering from the catastrophic loss of a major telecommunications office," *IEEE Commun. Mag.*, vol. 49, no. 1, pp. 28–35, Jan. 2011.

[15] T. Adachi, Y. Ishiyama, Y. Asakura, and K. Nakamura, "The restoration of telecom power damages by the Great East Japan Earthquake," in *IEEE 33rd International Telecommunications Energy Conference*, 2011, pp. 1–5.

[16] "Flooding, power outages from hurricane sandy lead to internet, phone service disruptions," http://nypost.com/2012/10/30/flooding-power-outages-from-hurricane-sandy-lead-to-internet-phone-service-disruptions, 2012.

[17] A. Kwasinski, "Lessons from field damage assessments about communication networks power supply and infrastructure performance during natural disasters with a focus on hurricane sandy," in *FCC Workshop Network Resiliency*, 2013.

[18] "2008sichuan earthquake," http://en.wikipedia.org/wiki/2008 Sichuan earthquake.

[19] "2011tohoku earthquake and tsunami," http://en.wikipedia.org/wiki/2011 Tohoku earthquake and tsunami.

[20] "2014 U.S. geological national seismic hazard maps," http://earthquake.usgs.gov/hazards/products/conterminous/index.php, 2014.

[21] "National hurricane center," http://www.nhc.noaa.gov.

[22] H. Nakamura, S. Horiuchi, C. Wu, S. Yamamoto, and P. A. Rydelek, "Evaluation of the real-time earthquake information system in japan," *Geophysical Research Letters*, vol. 36, no. L00B01, Jan. 2009.

[23] Reuters, "Experts warn of substantial risk of WMD attack," http://research.lifeboat.com/lugar.htm.

[24] B. Graham *et al.*, *World at Risk: The Report of the Commission on the Prevention of WMD Proliferation and Terrorism.* Vintage Books, A Division of Random House, Inc., 2008.

[25] J. P. G. Sterbenz *et al.*, "Resilience and survivability in communication networks: Strategies, principles, and survey of disciplines," *Computer Networks*, vol. 54, no. 8, pp. 1245–1265, Jun. 2010.

[26] B. Mukherjee, M. F. Habib, and F. Dikbiyik, "Network adaptability from disaster disruptions and cascading failures," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 230–238, May. 2014.

[27] R. S. Couto, S. Secci, M. M. Campista, and L. M. K. Costa, “Network design requirements for disaster resilience in IaaS clouds,” *IEEE Commun. Mag.*, vol. 52, no. 10, pp. 52–58, Oct. 2014.

[28] C. Doerr and F. A. Kuipers, “All quiet on the Internet front?” *IEEE Commun. Mag.*, vol. 52, no. 10, pp. 46–51, Oct. 2014.

[29] S. S. Savas, M. F. Habib, M. Tornatore, F. Dikbiyik, and B. Mukherjee, “Network adaptability to disaster disruptions by exploiting degraded-service tolerance,” *IEEE Commun. Mag.*, vol. 52, no. 12, pp. 58–65, Dec. 2014.

[30] I. B. B. Harter, D. A. Schupke, M. Hoffmann, and G. Carle, “Network virtualization for disaster resilience of cloud services,” *IEEE Commun. Mag.*, vol. 52, no. 12, pp. 88–95, Dec. 2014.

[31] S. Neumayer and E. Modiano, “Network reliability with geographically correlated failures,” in *IEEE INFOCOM*, 2010, pp. 1–9.

[32] S. Neumayer, G. Zussman, R. Cohen, and E. Modiano, “Assessing the vulnerability of the fiber infrastructure to disasters,” *IEEE/ACM Trans. Netw.*, vol. 19, no. 6, pp. 1610–1623, Dec. 2011.

[33] J. Liu, X. Jiang, H. Nishiyama, and N. Kato, “Reliability assessment for wireless mesh networks under probabilistic region failure model,” *IEEE Trans. Veh. Technol.*, vol. 60, no. 5, pp. 2253–2264, Jun. 2011.

[34] X. Wang, X. Jiang, and A. Pattavina, “Assessing network vulnerability under probabilistic region failure model,” in *IEEE 12th International Conference on High Performance Switching and Routing*, 2011, pp. 164–170.

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

[36] J. Xiao, H. Wen, B. Wu, X. Jiang, P.-H. Ho, and L. Zhang, “Joint design on DCN placement and survivable cloud service provision over all-optical mesh networks,” *IEEE Trans. Commun.*, vol. 62, no. 1, pp. 235–245, Jan. 2014.

[37] M. F. Habib, M. Tornatore, M. D. Leenheer, F. Dikbiyik, and B. Mukherjee, “A disaster-resilient multi-content optical datacenter network architecture,” in *2011 13th International Conference on Transparent Optical Networks*, 2011, pp. 1–4.

[38] ——, “Design of disaster-resilient optical datacenter networks,” *J. Lightw. Technol.*, vol. 30, no. 16, pp. 2563–2573, Aug. 2012.

[39] S. Ferdousi, F. Dikbiyik, M. F. Habib, and B. Mukherjee, “Disaster-aware datacenter and content placement in cloud networks,” in *2013 IEEE International Conference on Advance Networks and Telecommuncations Systems*, 2013, pp. 1–3.

[40] S. Ferdousi, F. Dikbiyik, M. F. Habib, M. Tornatore, and B. Mukherjee, “Disaster-aware datacenter placement and dynamic content management in cloud networks,” *IEEE/OSA Journal of Optical Communications and Networking*, vol. 7, no. 7, pp. 681–694, Jul. 2015.

[41] J. Yao, P. Lu, L. Gong, and Z. Zhu, “On fast and coordinated data backup in geo-distributed optical inter-datacenter networks,” *J. Lightw. Technol.*, vol. 33, no. 14, pp. 3005–3015, Jul. 2015.

[42] P. Lu, L. Zhang, X. Liu, J. Yao, and Z. Zhu, “Highly efficient data migration and backup for big data applications in elastic optical inter-data-center networks,” *IEEE Network*, vol. 29, no. 5, pp. 36–42, Sept.-Oct. 2015.

[43] S. Ferdousi, M. Tornatore, M. F. Habib, and B. Mukherjee, “Rapid data evacuation for large-scale disasters in optical cloud networks,” *IEEE/OSA Journal of Optical Communications and Networking*, vol. 7, no. 12, pp. B163–B172, Dec. 2015.

[44] P. Lu, Q. Ling, and Z. Zhu, “Maximizing utility of time-constrained emergency backup in inter-datacenter networks,” *IEEE Commun. Lett.*, vol. 20, no. 5, pp. 890–893, May 2016.

[45] M. F. Habib, M. Tornatore, F. Dikbiyik, and B. Mukherjee, “Disaster survivability in optical communication networks,” *Computer Communications*, vol. 36, no. 6, pp. 630–644, Mar. 2013.

[46] X. Dong, T. EI-Gorashi, and J. Elmirghani, “Green IP over WDM networks with data centers,” *J. Lightw. Technol.*, vol. 29, no. 12, pp. 1861–1880, Jun. 2011.

[47] J. Xiao, B. Wu, X. Jiang, A. Pattavina, H. Wen, and L. Zhang, “Scalable data center network architecture with distributed placement of optical switches and racks,” *IEEE/OSA Journal of Optical Communications and Networking*, vol. 6, no. 3, pp. 270–281, Mar. 2014.

[48] D. Applegate, A. Archer, V. Gopalakrishnan, S. Lee, and K. K. Ramakrishnan, “Optimal content placement for a large-scale VoD system,” *IEEE/ACM Trans. Netw.*, vol. 24, no. 4, pp. 2114–2127, Aug. 2016.

[49] K. L. Yeung and T.-S. P. Yum, “Node placement optimization in shufflenets,” *IEEE/ACM Trans. Netw.*, vol. 6, no. 3, pp. 319–324, Jun. 1998.

[50] P. Cheng, C.-N. Chuah, and X. Liu, “Energy-aware node placement in wireless sensor networks,” in *IEEE GLOBECOM*, vol. 5, 2004, pp. 3210–3214.

[51] B. Wang, H. Xu, W. Liu, and H. Liang, “A novel node placement for long belt coverage in wireless networks,” *IEEE Trans. Comput.*, vol. 62, no. 12, pp. 2341–3353, Dec. 2013.

[52] T. Gomes *et al.*, "A survey of strategies for communication networks to protect against large-scale natural disasters," in *2016 8th International Workshop on Resilient Networks Design and Modeling (RNDM)*, 2016, pp. 11–22.

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

[54] S. S. Savas, M. F. Habib, F. Dikbiyik, M. Tornatore, and B. Mukherjee, "Disaster-aware service provisioning with manycasting in could networks," *Photon Network Communications*, vol. 28, no. 2, pp. 123–134, Oct. 2014.

[55] J. Yao, P. Lu, and Z. Zhu, "Minimizing disaster backup window for geo-distributed multi-datacenter cloud systems," in *2014 IEEE International Conference on Communications (ICC)*, 2014, pp. 3631–3635.

[56] D. Boru, D. Kliazovich, F. Granelli, P. Bouvry, and A. Zomaya, "Energy-efficient data replication in cloud computing datacenters," *Cluster Computer*, vol. 18, pp. 385–402, Jan. 2015.

[57] R. S. Couto, S. Secci, M. E. M. Campista, and L. H. M. Costa, "Server placement with shared backups for disaster-resilient clouds," *Computer Networks*, vol. 93, pp. 423–434, Dec. 2015.

[58] N. H. Bao, M. F. Habib, M. Tornatore, C. U. Martel, and B. Mukherjee, "Global versus essential post-disaster re-provisioning in telecom mesh networks," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 7, no. 5, pp. 392–400, May 2015.

[59] C. N. da Silva, L. Wosinska, S. Spadaro, J. C. W. A. Costa, C. R. L. Francês, and P. Monti, "Restoration in optical cloud networks with relocation and services d-ifferentiation," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 8, no. 2, pp. 100–111, Feb. 2016.

[60] N.-H. Bao, M. Tornatore, C. U. Martel, and B. Mukherjee, "Fairness-aware degradation based multipath re-provisioning strategy for post-disaster telecom mesh networks," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 8, no. 6, pp. 441–450, Jun. 2016.

[61] J. Wang, C. Qiao, and H. Yu, "On progressive network recovery after a major disruption," in *IEEE INFOCOM*, 2011, pp. 1925–1933.

[62] S. Kamamura, D. Shimazaki, K. Genda, K. Sasayama, and Y. Uematsu, "Disaster recovery for transport network through multiple restoration stages," *IEICE Transactions on Communications*, vol. E98-B, no. 1, pp. 171–179, Jan. 2015.

[63] K. A. Sabeh, M. Tornatore, and F. Dikbiyik, "Progressive network recovery in optical core networks," in *2015 7th International Workshop on Reliable Networks Design and Modeling (RNDM)*, 2015, pp. 106–111.

[64] R. Goscien, K. Walkowiak, M. Klinkowski, and J. Rak, "Protection in elastic optical networks," *IEEE Network*, vol. 29, no. 6, pp. 88–96, Nov.-Dec. 2015.

[65] K. Miranda, A. Molinar, and T. Razafidralambo, "A survey on rapidly deployable solutions for post-disaster networks," *IEEE Commun. Mag.*, vol. 54, no. 4, pp. 117–123, Apr. 2016.

[66] P. N. Tran and H. Saito, "Geographical route design of physical networks using earthquake risk information," *IEEE Commun. Mag.*, vol. 54, no. 7, pp. 131–137, Jul. 2016.

[67] A. Mauthe *et al.*, "Disaster-resilient communication networks: principles and best practices," in *2016 8th International Workshop on Resilient Networks Design and Modeling (RNDM)*, 2016, pp. 1–10.

[68] M. Tornatore *et al.*, "A survey on network resiliency methodologies against weather-based disruptions," in *2016 8th International Workshop on Resilient Networks Design and Modeling (RNDM)*, 2016, pp. 23–34.

[69] "2014 U.S. geological national seismic hazard map gridded data," http://earthquake.usgs.gov/hazards/products/conterminous/2014/data/2014 pga2pct50yrs.dat.zip, 2014.

[70] "InternetMCI network," http://www.topology-zoo.org/dataset.html, 2011.

Publications

Journal Articles

[1] Lisheng Ma, Xiaohong Jiang, Bin Wu, Achille Pattavina and Norio Shiratori. Probabilistic region failure-aware data center network and content placement. *Computer Networks*, vol. 103, pp. 56-66, July 2016.

[2] Lisheng Ma, Xiaohong Jiang, Bin Wu and Achille Pattavina. Performance analysis of f-cast crosstalk-free optical banyan networks. *IEEE Transactions on Communications*, vol. 65, no. 4, pp. 1721-1732, April 2017.

[3] Lisheng Ma, Wei Su, Bin Wu, Tarik Taleb, Xiaohong Jiang and Norio Shiratori. ε-Time early warning data backup in disaster-aware optical inter-connected data center networks. *IEEE/OSA Journal of Optical Communications and Networking*, vol. 9, no. 6, pp. 536-545, June 2017.

[4] Lisheng Ma, Bin Wu, Xiaohong Jiang and Achille Pattavina. Nonblocking conditions for (f_1, f_2)-cast Clos networks under balanced traffic. *Optical Switching and Networking*, vol. 25, pp. 109-116, July 2017.

Conference Papers

[5] Lisheng Ma, Xiaohong Jiang, Achille Pattavina and Norio Shiratori Probabilistic region failure-aware data center network placement. IEEE 16th International Conference on High Performance Switching and Routing (HPSR), 2015, pp. 1-6.

[6] Lisheng Ma, Xiaohong Jiang, Bin Wu, Tarik Taleb, Achille Pattavina and Norio Shiratori. Cost-efficient data backup for data center networks against ε-time early warning disaster. IEEE 17th International Conference on High Performance Switching and Routing (HPSR), 2016, pp. 22-26.