

# Evaluation of Topological Vulnerability of the Internet under Regional Failures

Wei Peng, Zimu Li, Jinshu Su, Muwei Dong

► **To cite this version:**

Wei Peng, Zimu Li, Jinshu Su, Muwei Dong. Evaluation of Topological Vulnerability of the Internet under Regional Failures. A Min Tjoa; Gerald Quirchmayr; Ilsun You; Lida Xu. 1st Availability, Reliability and Security (CD-ARES), Aug 2011, Vienna, Austria. Springer, Lecture Notes in Computer Science, LNCS-6908, pp.164-175, 2011, Availability, Reliability and Security for Business, Enterprise and Health Information Systems. <10.1007/978-3-642-23300-5\_13>. <hal-01590391>

**HAL Id: hal-01590391**

**<https://hal.inria.fr/hal-01590391>**

Submitted on 19 Sep 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# Evaluation of Topological Vulnerability of the Internet under Regional Failures

Wei Peng<sup>1</sup>, Zimu Li<sup>2</sup>, Jinshu Su<sup>1</sup>, and Muwei Dong<sup>1</sup>

<sup>1</sup> School of Computer, National University of Defense Technology,  
Changsha, Hunan, 410073, China  
{wpeng, sjs, mwdong}@nudt.edu.cn

<sup>2</sup> Network Research Center, Tsinghua University, Beijing, 100084, China  
lzm@cernet.edu.cn

**Abstract.** Natural disasters often lead to regional failures which can fail down network nodes and links co-located in a large geographical area. It will be beneficial to improve the resilience of a network by assessing its vulnerability under regional failures. In this paper, we propose the concept of  $\alpha$ -critical-distance to evaluate the importance of a network node in the geographical space with a given failure impact ratio  $\alpha$ . Theoretical analysis and a polynomial time algorithm to find the minimal  $\alpha$ -critical-distance of a network are presented. Using real Internet topology data, we conduct experiments to compute the minimal  $\alpha$ -critical-distances for different networks. The computational results demonstrate the differences of vulnerability of different networks. We also find that with the same impact ratio  $\alpha$ , the studied topologies have smaller  $\alpha$ -critical-distances when the network performance is measured by network efficiency than giant component size.

**Keywords:** network topology, vulnerability, regional failure, critical distance, algorithm

## 1 Introduction

Complex networks like the Internet play an important role in today's world. The society and our lives heavily depend on these networks. For example, most of data communications are accomplished through the Internet. However, as a fast-growing system, the Internet may become vulnerable to intrusions and failures since it integrates more and more components, services and functions. It is generally believed that the current Internet is not so robust or resilient to failures as a critical infrastructure. A breakdown of the Internet for a short while may result in the loss of a large fortune. Considering the importance and the central role of the Internet, its vulnerability has been studied from various aspects.

The topology of a network ranks foremost among all factors which influence the network vulnerability. During the past years, great efforts have been paid in evaluating the vulnerability of a network topology and finding efficient algorithms to enhance the robustness of a network. In most of the works, a network

topology is studied as a pure graph without considering the geographical properties of nodes and links. The network vulnerability is often studied on the logical network topology instead of the physical network topology.

Network nodes or links in a large geographical area may fail to work during the outbreak of electric black-outs and natural disasters like earthquakes, floods and tornadoes. For example, the Taiwan Earthquake in December 2006 damages several undersea cable systems in Asia and the communications between many Asia sites and U.S. sites are disrupted for several ten minutes to hours later. Such regional failures have geographical locations and exhibit strong space correlation among network nodes and links. Therefore, the geographical distribution of a network will have great impact on its vulnerability on regional failures. It has been shown that complex networks including the Internet are resilient to independent random failures, but fragile to intentional attacks [1]. The vulnerability of the Internet under regional failures is still an open issue.

In this paper, we study the vulnerability of the Internet under large-scale regional failures using its city-level topology. A large-scale regional failure means that a significant number of network nodes or links located in a geographical area become unavailable due to natural disasters, electrical black-outs or terroristic attacks. By studying the vulnerability of the Internet in geographical space, we can understand the impact of the geographical properties of the Internet on regional failures. It also helps to identify not only the critical nodes and links, but also the critical area in the network. By the study, we will evaluate the importance of network nodes under regional failures and explore the necessary failure impact range to incur a certain level of degradation of network performance.

To achieve the goal, we develop a binary disk-shape regional failure model as a start. Then the concept of  $\alpha$ -critical-distance is introduced where  $\alpha$  is a given failure impact ratio. The properties of the mapping function from failure impact range to failure impact ratio is analyzed. Based on the properties, a polynomial time algorithm is proposed to calculate the minimal  $\alpha$ -critical-distance given the impact ratio  $\alpha$  and a network topology. We collect topology data sets from the Internet and annotate network nodes with geographical locations. Using the real topology data, we evaluate and compare the network vulnerability of several countries. The experimental results reveal the differences among the studied topologies. We also find that smaller failure impact distances are required when the network efficiency [2] is used as the network performance metric than the giant component size.

The rest of the paper is organized as follows. Section 2 presents a short survey on the related work. Section 3 describes the vulnerability model and the algorithm to find the minimal  $\alpha$ -critical-distance. The experimental results using real topology data are presented in section 5. We conclude the paper at last.

## 2 Related Work

Numerous efforts have been devoted to the study of the network vulnerability, including proposing vulnerability measures and methods to evaluate the vulner-

ability of a network, proposing measures and methods to find critical nodes and links in a network. However, few works have explored the network vulnerability issues under geographical correlated failures.

Basic network vulnerability measures are derived from global graph properties like the number of nodes and edges, the lengths of pairwise shortest paths. These measurement criteria include average or characteristic path length, giant component size, network efficiency [2], N-Q measure [3], algebraic connectivity and natural connectivity [4]. Based on these measures, the vulnerabilities of networks with different characteristics are investigated. For example, the giant component size is widely used in the research of complex networks. It is shown that scale-free complex networks including the Internet is vulnerable to intentional attacks with their power-law features [1]. The algebraic connectivity is used to study the robustness of networks of three types subject to random node and link failures [5]. Dinh et al. [6] investigated the measure called *pairwise connectivity* and formulate the vulnerability assessment problem as a *NP-complete* problem and present pseudo-approximation algorithms to detect the set of critical nodes and critical edges.

To assess the importance of nodes on the vulnerability of a network, local nodal measures are proposed, including degree centrality, betweenness centrality and closeness centrality [7]. In [8], a distributed algorithm for locating critical nodes is proposed based on spectral analysis. Kermerrec et al. [9] propose a measurement called *second order centrality* and use random walk to assess node criticality in a distributed manner. Arulsevan et al. [10] have shown that the *critical node problem* in a graph is a *NP-complete* problem and proposed a heuristic algorithm for it. Other approaches consider logical network attacks [11] and the role of humans in system security models [12].

All above approaches aim at the logical network topology, without considering the geographical properties and correlated regional failure effects.

Hansen et al. [13] studied the network resilience under regional failures at routing layer. Grubestic et al. [14] Studied the *Node Removal Impact Problem* in the geographical context. In [15], the vulnerability of networks is assessed in geographical space by modeling the physical network as a bipartite graph. The location of geographical disasters that maximize the capacity of disconnected links is studied. Specifically, a polynomial time algorithm for finding the location of a vertical line segment cut is proposed. The problem is further studied by applying geometric probability techniques to calculate network performance measures on a random line-cut [16]. Agarwal et al. [17] establish a probabilistic geographical failure model and use computational geometric tools to identify vulnerable points within a network. Their algorithms are applied in three optical WDM networks within the continental USA. In [18], both nodes and links located in failure impact area are assumed to be down after a regional failure. Using the European network COST-239 as test network, the authors present some preliminary results by simulating regional failures with three impact ranges.

The previous works try to find the optimal location of geographical failures that maximize the failure impact with a given failure impact range. In contrast

with these works, we try to find the optimal regional failure area or the weakest portion of a network that decreases the network performance to a threshold value.

### 3 Vulnerability Evaluation Method

#### 3.1 Regional Failure Model

Regional failures belong to the category of multiple-failure scenarios and a large-scale regional failure often occurs in a geographical area where nodes or links may crash or fail to work. A regional failure can affect only a small part of the network, but it can also damage the network severely if it happens in a crucial area and affects enough number of network nodes or links. The area where a regional failure takes effects is defined as the failure area. A regional failure is described by its failure area and the failure probability of nodes and links in the area.

The failure area of a regional failure can be in any shapes. For simplicity, we define a *binary disk-shape failure model* which is characterized by its epicenter position and the failure impact range in a 2-dimension space.

**Definition 1:** The binary disk-shape failure model is defined as  $(L0, r)$ , where  $L0$  is the location of the failure center and  $r$  is the failure impact range. For node  $i$ , if its distance to the failure center is measured by  $dd_i$ , the state of node  $i$  after the failure will be:

$$s_i = \begin{cases} 0 & , \quad dd_i \leq r \\ 1 & , \quad else \end{cases} \quad (1)$$

The state 0 means that the node is down while the state 1 represents that the node is still up.

The failure model defined above is deterministic, while one could use more complex models like the probability failure model in [17]. In many cases, a simple model with less complexity can provide enough information for network operators, so we focus on the binary disk-shape failure model in this paper. Our model is similar to that in [18] and the difference is that we do not consider links traversing the failure area.

#### 3.2 Vulnerability Measurement

Many statistical metrics have been proposed to quantify the failure impact. In this paper, we consider the following performance metrics:

- $S$ : giant component size, which is the number of nodes in the largest connected sub-graph after a failure;
- $E$ : network efficiency [2], which is defined as the sum of inverse values of all shortest path lengths in a graph. Given a graph  $G$ , its network efficiency is defined as

$$E(G) = \frac{1}{n(n-1)} \sum_{i \neq j \in G} \frac{1}{d_{i,j}} \quad (2)$$

where  $n$  is the number of nodes in  $G$  and  $d_{i,j}$  is the shortest path length from node  $i$  to node  $j$ .

The giant component size is widely used in assessing the robustness of complex networks under random failures and intentional attacks. It can characterize network partitioning to some extent. However, if the network remains connected after a failure, it will fail to quantify the impact of the failure. Comparatively, the change of network topology can be revealed by the network efficiency since it will decrease with the increasing of path lengths after a failure. So it is better to assess the performance degradation using network efficiency than giant component size if the computational complexity is bearable.

Given the parameters of a failure and a network topology, we can calculate the performance metrics before and after the failure. The evaluation of the vulnerability of a network is the reverse problem. That is, we need to find the minimal failure impact range to break down a network and the crucial portion of the network. The weakest area in a network is the Achilles' heels of the network.

To achieve the goal, we introduce the concept of  $\alpha$ -critical-distance which is the minimal impact range in a regional failure when the network performance metric is reduced to  $\alpha$  times of the initial value before the failure.  $\alpha$  ( $0 \leq \alpha \leq 1$ ) is called the *impact ratio*. Like the approach in [18], we assume that the epicenter is the location of one node in the network. For node  $i$ , we use the mapping  $F^i$  to define the relationship between the impact ratio  $\alpha$  and the impact range  $r$ . Formally,

$$\alpha = F^i(r) \quad (3)$$

The reverse mapping of  $F$  is denoted as  $F^{i(-1)}$ .

**Definition 2:** Given  $\alpha$ , the  $\alpha$ -critical-distance of node  $i$  is the minimal impact range  $r_i^\alpha$  when node  $i$  locates at the epicenter. That is,

$$r_i^\alpha = \min F^{i(-1)}(\alpha) \quad (4)$$

The mapping  $F^i$  of node  $i$  has the following properties.

**Property 1:** For any  $r > 0$ ,  $0 \leq F^i(r) \leq 1$ .

**Property 2 (Monotonicity):** If  $r_1 > r_2$ , then  $F^i(r_1) \leq F^i(r_2)$ .

**Property 3:** If the maximal distance from node  $i$  to any other nodes in the network is  $d_i^{max}$ , then for any  $r \geq d_i^{max}$ ,  $F^i(r) = 0$ .

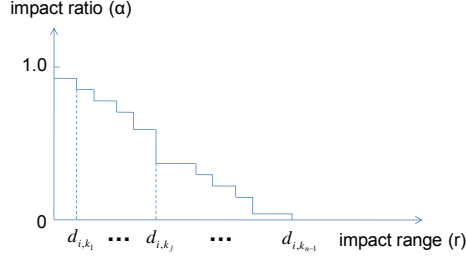
The property 3 means that all nodes will go down if the failure impact range covers all nodes in the network. The network performance is measured as zero when all nodes go down or isolated and the network is totally unavailable.

Let the distance from node  $i$  to node  $j$  be  $d_{i,j}$ ,  $1 \leq j \leq n$ ,  $j \neq i$ . We sort the distances from node  $i$  to other nodes in ascending order. Let the sorted distances be  $d_{i,k_1} \leq d_{i,k_2} \leq \dots \leq d_{i,k_{n-1}}$ . And we assume that nodes co-located with node  $i$  on the same position will be down when the failure impact range  $r = 0$ .

**Property 4:** (1) If  $d_{i,k_j} \leq r < d_{i,k_{j+1}}$ ,  $1 \leq j \leq (n-2)$ , then  $F^i(r) = F^i(d_{i,k_j})$ ;

(2) if  $0 \leq r < d_{i,k_1}$ , then  $F^i(r) = F^i(0)$ ;

(3) if  $r \geq d_{i,k_{n-1}}$ , then  $F^i(r) = F^i(d_{i,k_{n-1}}) = 0$ .



**Fig. 1.** Illustration of Property 4 and 5

Based on the property 4, we can get the conclusion that there will be at most  $n$  distinct values of  $\alpha$  for each  $F^i$ .

**Property 5:** if  $p > q$ ,  $F^i(d_{i,k_p}) \leq F^i(d_{i,k_q})$ .

Figure 1 is an example to illustrate the Property 4 and 5. In fact, the mapping  $F^i$  is a monotonic step-decreasing function of failure impact range  $r$ .

For node  $i$ , let  $F_p^i = F^i(d_{i,k_p})$  for short. Based on the sorted distances  $(d_{i,k_1}, d_{i,k_2}, \dots, d_{i,k_{n-1}})$ , we define  $E_p^i$  as the sequence of node  $i$  followed by its  $p$  nearest nodes.

**Definition 3:** The  $p$ -node-sequence of node  $i$  is defined as  $E_p^i = (i, k_1, k_2, \dots, k_p)$ . The complementary  $p$ -node-sequence is defined as  $CE_p^i = (k_{p+1}, \dots, k_{n-1})$ .

If a node  $v$  is in the  $p$ -node-sequence, it is denoted as  $v \in E_p^i$ . If for any  $v \in E_p^i$  and any  $u \in E_p^j$ ,  $u \in E_p^i$  and  $v \in E_p^j$ , then we say  $E_p^i = E_p^j$ . In such a case, the regional failure centered at node  $i$  with the impact range  $d_{i,k_p}$  will take the same effects on the network as the failure centered at node  $j$  with the impact range  $d_{j,k_p}$ .

**Property 6:** For any  $p$ ,  $1 \leq p \leq (n-1)$ , if  $E_p^i = E_p^j$ , then  $F_p^i = F_p^j$ .

**Property 7:** If the complementary  $p$ -node-sequence  $CE_p^i$  and  $CE_p^j$  have the same node sequence, then for any  $m$ ,  $(p+1) \leq m \leq (n-1)$ ,  $F_m^i = F_m^j$ .

**Property 8:** For any node  $i$  and  $j$ ,  $\max\{0, (r_j^\alpha - d_{i,j})\} \leq r_i^\alpha \leq (r_j^\alpha + d_{i,j})$  and  $\max\{0, (r_i^\alpha - d_{i,j})\} \leq r_j^\alpha \leq (r_i^\alpha + d_{i,j})$ .

**Property 9:** For any node  $i$ ,  $\max_{1 \leq j \leq n} \{0, (r_j^\alpha - d_{i,j})\} \leq r_i^\alpha \leq \min_{1 \leq j \leq n} (r_j^\alpha + d_{i,j})$ .

Property 7 is a special case of Property 6, because if  $CE_p^i$  and  $CE_p^j$  have the same node sequence, then for any  $m$ ,  $(p+1) \leq m \leq (n-1)$ ,  $E_m^i = E_m^j$ . Property 8 relates the  $\alpha$ -critical-distances of two nodes with the distance between them. Property 9 can be induced from Property 8 directly.

For different values of  $\alpha$ , a node will have different  $\alpha$ -critical-distance values. Given the value of  $\alpha$ , we can assess the importance of the location of a node with its  $\alpha$ -critical-distance. Smaller the  $\alpha$ -critical-distance is, more important the node location is. Therefore, we evaluate the vulnerability of a network by finding the node with the smallest  $\alpha$ -critical-distance in the network.

### 3.3 Finding Minimal $\alpha$ -critical-distance

Under regional failure scenarios, the node having the minimal  $\alpha$ -critical-distance is called the *critical node* of the network. The problem of finding the critical node of a network can be described as:

$$\begin{aligned} &\text{Given a network } G \text{ and } \alpha, \text{ find } i \\ &\text{s.t. } r_i^\alpha = \min \{r_j^\alpha, 1 \leq j \leq n\} \end{aligned}$$

Because a node has  $n$  distinct values of  $\alpha$  at most, the  $\alpha$ -critical-distance of node  $i$  can be computed by enumerating all  $n$  possible cases. The time complexity of computing  $r_j^\alpha$  will be  $O(n * T)$  and the worst time complexity of the problem would be  $O(n^2 * T)$ . Here  $T$  is the time complexity to calculate the mapping  $F^i(r)$  with given failure impact range  $r$ . The time to compute the giant component size is  $O(m)$  where  $m$  is the number of edges in the graph. The time to compute the network efficiency is  $O(n^3)$ . Thus, the worst time complexity of the problem would be  $O(mn^2)$  and  $O(n^5)$ , respectively.

**Table 1.** Variables of Algorithm 1

Input Variable	Description
$G$	network represented by an undirected graph
$n$	the number of nodes in $G$
$\alpha$	impact ratio of the network performance
$K$	the number of minimal $\alpha$ -critical-distances
Output Variable	Description
$M$	the vector of $\alpha$ -critical-distances, $M[i]$ is the $i$ -th $\alpha$ -critical-distance
Temporary Variable	Description
$d_{i,j}$	geometric distance between node $i$ and node $j$
$d_{min}$	minimal value of $\alpha$ -critical-distance of a node
$d_{max}$	maximal value of $\alpha$ -critical-distance of a node
$rset$	the set of candidate values of $\alpha$ -critical-distance

For large-scale networks, it will be time-costly to find critical nodes in a network. Because the  $\alpha$ -critical-distance of a node is related to others according to Property 8, we can use the results computed in previous steps to reduce the time complexity. Here we propose a fast algorithm to find  $K$  minimal  $\alpha$ -critical-distance values. The input and output variables are summarized in Table 1.

The algorithm calculates the  $\alpha$ -critical-distance for each node iteratively. In step 2 and step 3, the algorithm finds the lower bound and the upper bound of  $\alpha$ -critical-distance that node  $i$  can take. According to Property 9, the lower and upper bounds can be computed with the  $\alpha$ -critical-distances of other nodes found in previous steps. Then the set of candidate values of  $\alpha$ -critical-distance is generated by selecting nodes which distances fall into the bounds. The procedure *sortCandidateDistances(rset)* sorts the candidate distance set  $rset$  in ascending



order. The procedure  $binarySearch(G, i, rset)$  finds the  $\alpha$ -critical-distance of node  $i$  by applying the *binary search* method on the candidate distance set  $rset$ . After the  $\alpha$ -critical-distance values of all nodes are found, the procedure  $sortCriticalDistances(M, K)$  sorts the values in ascending order and outputs the first  $K$  values.

By using the *binary search* method, the time complexity to find the  $\alpha$ -critical-distance of a node can be reduced from  $O(n*T)$  to  $O(\log_2 n * T)$ . Both the step 8 and step 11 have the worst-case time complexity of  $O(n^2)$ , e.g., if the Quicksort algorithm is applied. Thereby, the time complexity of the algorithm will be  $O(mn \log_2 n + n^3 + n^2)$  and  $O(n^4 \log_2 n + n^3 + n^2)$ , if the giant component size and the network efficiency are used as performance metrics respectively.

---



---

**Algorithm 1:** Find  $K$  minimal  $\alpha$ -critical-distances

---

```

1: for  $i = 1$  to  $n$  do
2:    $d_{min} = \max_{1 \leq j \leq (i-1)} \{0, M[j] - d_{i,j}\};$ 
3:    $d_{max} = \min_{1 \leq j \leq (i-1)} \{M[j] + d_{i,j}\};$ 
4:    $rset = \emptyset;$ 
5:   for  $j = 1$  to  $n$  do
6:     if  $d_{min} \leq d_{i,j} \leq d_{max}$  then
7:        $rset = rset \cup \{d_{i,j}\};$ 
8:     end for
9:    $sortCandidateDistances(rset);$ 
10:   $M[i] = binarySearch(G, i, rset);$ 
11: end for
12:  $sortCriticalDistances(M, K);$ 

```

---

## 4 Computational Results

### 4.1 Data Sets

To study the vulnerability of the Internet topology and evaluate the algorithm proposed in the above section, we collect some online data sets. Since an autonomous system (AS) may span across a large geographical area, the AS-level Internet topology is not appropriate for the purpose of this study. The router-level Internet topology is ideal for the study. However, there are few available data sets and a lot of links are missing in the public data sets due to privacy issues and limitations of network measurement methods. Instead, we use the city-level Internet topology in this paper.

The data sets come from DIMES [19]. In the city-level topology data sets given by DIMES, the data set of July 2009 has the largest number of nodes and links. We extract the city-level topology from it for several countries, including United Kingdom (UK), Japan (JP), India (IN) and China (CN). The location information for each city in the original data of DIMES lacks precision, so we assign each city with the longitude and latitude looked up from the data provided by MaxMind [20]. Because many cities in the original topology have the

same coordinates, we merge them into one city and make the resultant city-level topology compact. We use abbreviation of the country name to name the topology data set for each country. The basic attributes of the topology data sets are shown in Table 2. It should be noted that the topology data set is just a snapshot of the Internet’s dynamic structure and the degree of incompleteness may vary greatly for different countries.

**Table 2.** Basic Attributes of Topology Data Sets

Data Set	Node Number	Edge Number	Average Degree	Average Path Length
UK	1085	4044	7.454378	2.371940
JP	117	856	14.632479	1.915119
IN	349	1101	6.309456	2.225933
CN	922	2002	4.342733	2.900891

#### 4.2 Evaluation of the Proposed Algorithm

We evaluate the proposed algorithm at first. The giant component size is used as the performance metric. To demonstrate the efficiency of the optimization operations using distance bounds (step 2-9 in algorithm 1), we compare the algorithm with and without the optimization steps. If the optimization steps are not applied, all distance values will be used in step 10, e.g.,  $rset = \{d_{i,j}, 1 \leq j \leq n\}$ . Since computing the impact ratio with a given failure area takes the longest time, we use the number of function calls to evaluate the algorithm performance. The *performance improved ratio* (PIR) is defined as:

$$PIR = \frac{C^{orig} - C^{opt}}{C^{orig}} \quad (5)$$

where  $C^{orig}$  is the number of function calls without optimization steps and  $C^{opt}$  is that with optimization.

Figure 2 demonstrates the performance of the algorithm. It is shown that the algorithm performance can be improved with the optimization steps for different network topologies. Normally, about half of the function calls can be saved, while 45% function calls can be saved in the worst case. Besides, when the impact ratio ( $\alpha$ ) is approaching its extreme value (0 or 1), the algorithm have better performance than the case when  $\alpha$  is in the middle of its value range.

#### 4.3 Evaluation of the Topologies

We use the proposed algorithm to find the  $\alpha$ -critical-distances for different values of  $\alpha$ . Figure 3 shows that with the increasing of  $\alpha$ , the  $\alpha$ -critical-distance

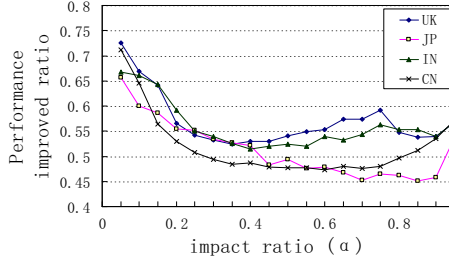


Fig. 2. Improved ratio with optimization

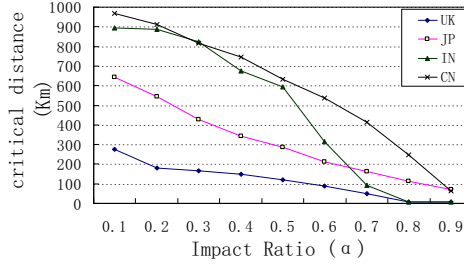


Fig. 3.  $\alpha$ -critical-distance vs Impact Ratio

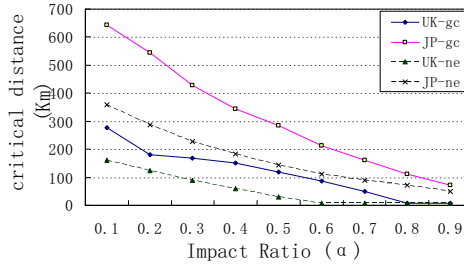


Fig. 4.  $\alpha$ -critical-distance with Different Metrics

decreases. For different countries, different failure impact ranges are required to get the same impact ratio. When  $\alpha$  is less than 0.6, India and China have the greater  $\alpha$ -critical-distance values than UK and Japan. It is reasonable since large failure impact ranges are required for countries with large acreage. Japan has longer critical distances than UK for the same impact ratio. Since the topology snapshot of Japan has less nodes but higher node degrees than the topology of UK, the former is more resilient to disk-shape regional failures than the latter.

Table 3 shows the top 10 minimal  $\alpha$ -critical-distances when  $\alpha=0.5$ . It is shown that a regional failure with a radius of 120 Kilometers in UK may lead to half of network nodes disconnected to others.

**Table 3.** Top 10 Critical Distances (Km) ( $\alpha=0.5$ )

rank	UK	Japan	India	China
1	119.014675	284.910063	594.922675	633.075876
2	119.082917	286.358551	596.185520	634.302529
3	119.086647	287.783538	605.273073	640.487066
4	119.300718	291.704459	606.387328	645.156163
5	119.467601	293.769890	606.453596	646.118720
6	119.510586	301.692601	607.552635	647.432880
7	119.889240	301.952568	613.332279	647.609953
8	120.036688	313.328128	617.205720	649.372846
9	120.250490	317.550703	619.813512	649.941193
10	120.422221	319.867304	622.285994	652.425884

With network efficiency as the performance metric, the algorithm finds the  $\alpha$ -critical-distances for networks of UK and Japan. The results are shown in Figure 4. The suffix ‘-gc’ and ‘-ne’ are used to differentiate the curves using giant component size and network efficiency as the performance metrics respectively. Noteworthy differences can be observed. When the network efficiency is used, the network performance degrades even when regional failures have a small impact range. For example, in the worst case, a regional failure with an impact range of 29Km can lead to the network efficiency of UK’s network reduced to half of the normal performance. According to the definition of the network efficiency, it is more sensitive to the change of a network topology than the giant component size. Thus, in the context of regional failures, it is easy to harm the network efficiency of a network with small failure impact range.

## 5 Conclusion

In this paper, we introduce the concept of  $\alpha$ -critical-distance for the purpose of evaluating vulnerability of networks under regional failures. We define a binary disk-shape regional failure model and study the properties of the mapping from failure impact range to failure impact ratio. It is shown that the mapping is a monotonic step-decreasing function. Based on the theoretic analysis, we propose an algorithm to compute the  $\alpha$ -critical-distance for a given network topology and  $\alpha$ . Using real topology data sets, we run the algorithm to evaluate the vulnerability of city-level Internet topologies for several countries. The computational results reveal the differences of network vulnerabilities among different countries and the differences when different network performance metrics are used.

Our work provides a new way to assess the network vulnerability under large-scale regional failures. The future work includes studying the network vulnerability under regional failures which take effects both on nodes and edges, correlated analysis of critical distances and node densities, and so on.

**Acknowledgments.** The work is supported by the grants from the National Natural Science Foundation of China under grant No.61070199 and 60873214, and the National 973 Grand Fundamental Research Program 2009CB320503.

## References

1. Doyle J.C., Alderson D., Li L., Lowet S., et al.: The ‘Robust Yet Fragile’ Nature of the Internet. *PNAS*, 102 (41) (2005)
2. Latora V., Marchiori M.: Efficient Behavior of Small-World Networks. *Phys. Rev. Lett.*, Nov. 87:198701 (2001)
3. Nagurney A., Qiang Q.: A Network Efficiency Measure with Application to Critical Infrastructure Networks. *J. Glob. Optim.* 40, 261–275 (2008)
4. Wu J., Barahona M., Tan Y.-J., Deng H.-Z.: Natural Connectivity of Complex Networks. *Chinese Physical Lett.*, 27(7) 078902 (2010)
5. Jamakovic A., Uhlig S.: Influence of the network structure on robustness. In: 15th IEEE Int. Conf. on Networks (ICON), pp. 278–283, Adelaide, Australia (2007)
6. Dinh T.N., Xuan Y., Thai M.T., Pardalos P.M.: On New Approaches of Assessing Network Vulnerability: Hardness and Approximation. *IEEE/ACM Trans. on Networking*, under revision (2011)
7. Borgatti S.P., Everett M.G.: A graph-theoretic perspective on centrality. *Social Networks*, 28(4), 466–484 (2006)
8. Wehmuth K., Ziviani A.: Distributed Algorithm to Locate Critical Nodes to Network Robustness based on Spectral Analysis. *CoRR* abs/1101.5019 (2011)
9. Kermarrec A.-M., Merrer E.L., Sericola B., Trédan G.: Second order centrality: Distributed assessment of nodes criticality in complex networks. *Computer Communications*, 34, 619–628 (2011)
10. Arulselvan A., Commander C.W., Eleftheriadou L., Pardalos P.M.: Detecting Critical Nodes in Sparse Graphs. (2009)
11. Crescenzo G.D., Ghosh A., Kampasi A., Talpade R., Zhang Y.: Detecting Anomalies in Active Insider Stepping Stone Attacks. *J. Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, 2(1), 103–120 (2011)
12. Pieters W.: Representing Humans in System Security Models: An Actor-Network Approach. *J. Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, 2(1), 75–92 (2011)
13. Hansen A.F., Kvalbein A., Čičić T., Gjessing S.: Resilient Routing Layers for Network Disaster Planning. In: 4th Int. Conf. on Networking (ICN). LNCS, vol. 3421, pp. 1097–1105 (2005)
14. Grubestic T.H., Murray A.T.: Vital Nodes, Interconnected Infrastructures, and the Geographies of Network Survivability. *Annals of the Association of American Geographers*, 96(1), 64–83 (2006)
15. Neumayer S., Zussman G., Cohen R., Modiano E.: Assessing the Impact of Geographically Correlated Network Failures. In: *IEEE MILCOM 2008* (2008)
16. Neumayer S., Modiano E.: Network Reliability With Geographically Correlated Failures. In: *IEEE INFOCOM 2010*, pp. 1–9. San Diego, CA (2010)
17. Agarwal P.K., Efrat A., Ganjugunte S., Hay D., Sankararaman S., Zussman G.: The Resilience of WDM Networks to Probabilistic Geographical Failures. (2011)
18. Bassiri B., Heydari S. S.: Network Survivability in Large-Scale Regional Failure Scenarios. In: *C<sup>3</sup>S<sup>2</sup>E – 09*, Montreal, QC, Canada (2009)
19. The DIMES project, <http://www.netdimes.org/>
20. MaxMind GeoLite City, <http://www.maxmind.com/app/geolitecity/>