New Journal of Physics
The open-access journal for physics

Geographical constraints to range-based attacks on links in complex networks

Baihua Gong[1], **Jun Liu**[1], **Liang Huang**[1], **Kongqing Yang**[1,2] **and Lei Yang**[3,4]
[1] Department of Physics, Lanzhou University, Lanzhou 730000, People's Republic of China
[2] Institute of Applied Physics, Jimei University, Xiamen 361021, People's Republic of China
[3] Institute of Modern Physics, Chinese Academy of Science, Lanzhou 730000, People's Republic of China
E-mail: lyang@impcas.ac.cn

New Journal of Physics **10** (2008) 013030 (7pp)
Received 6 September 2007
Published 25 January 2008
Online at http://www.njp.org/
doi:10.1088/1367-2630/10/1/013030

Abstract. In this paper, we studied range-based attacks on links in geographically constrained scale-free networks and found that there is a continuous switching of roles of short- and long-range attacks on links when tuning the geographical constraint strength. Our results demonstrate that the geography has a significant impact on the network efficiency and security; thus one can adjust the geographical structure to optimize the robustness and the efficiency of the networks. We introduce a measurement of the impact of links on the efficiency of the network, and an effective attacking strategy is suggested.

Much attention has been directed to the study of small-world networks since Watts and Strogatz (WS) introduced their famous model [1]. This model is constructed from a sparse regular network by rewiring a small fraction of links at random. Watts [2] also introduced the concept of *range* to characterize different types of links: the *range* of a link l_{ij} connecting nodes i and j is defined as the length of the shortest path between nodes i and j in the absence of l_{ij}. Typically, local connections are short-range links but rewired connections are long-range links. The WS model is more sensitive to attacks on long-range links connecting nodes that would otherwise be separated by a long shortest path. This is not true for many scale-free networks. Motter *et al* [3] showed that short-range links, rather than long-range ones are the vital ones for efficient

[4] Author to whom any correspondence should be addressed.

1367-2630/07/013030+07$30.00
© IOP Publishing Ltd and Deutsche Physikalische Gesellschaft

communication between nodes in these networks. Gong *et al* [4] gave an analytical argument to this observation in the case of random scale-free (RSF) networks.

To study range-based attacks on links in complex networks, we consider the *efficiency* of the network, which is defined as [5]:

$$E = \frac{2}{N(N-1)} \sum \frac{1}{d_{ij}}, \tag{1}$$

where d_{ij} is the length of the shortest path between the node pair (i, j), the sum over all $N(N-1)/2$ pairs of nodes. The *efficiency* of a network is a measure of how efficiently it exchanges information, and it has a finite value even for a disconnected network. The fundamental difference between E and $1/D$ (D is the expected diameter of the network) is that E is the efficiency of a parallel system, where all the nodes in the network concurrently exchange packets of information (such as all the systems in [1], for example), while $1/D$ measures the efficiency of a sequential system (i.e. only one packet of information goes along the network) [5]. Consider the link l_{ij}, there are $L(i, j)$ ($L(i, j)$ is the *load* of l_{ij}) pairs of nodes whose shortest path passes l_{ij}. When removing l_{ij}, the distance between those node pairs has an increment of $R(i, j)-1$ on average ($R(i, j)$ is the *range* of l_{ij}), thus the decrement of the *efficiency* is approximately:

$$\begin{aligned}\Delta E &\approx \frac{2}{N(N-1)} \sum_{(m,n)} \left(\frac{1}{d_{mn}+R(i,j)-1} - \frac{1}{d_{mn}} \right) \\ &\approx \frac{2}{N(N-1)} \sum_{(m,n)} \frac{R(i,j)-1}{(d_{mn}+R(i,j)-1)d_{mn}} \\ &\sim \frac{2(R(i,j)-1)L(i,j)}{D(D-R(i,j))N(N-1)},\end{aligned} \tag{2}$$

where the sum is over all the node pairs whose shortest length increases as a result of the removal of l_{ij}. The *load* of a link is defined as the number of shortest paths passing through this link [6, 7]. When D is much greater than $R(i, j)$, the above equation can be rewritten as

$$\Delta E \sim \frac{2(R(i,j)-1)L(i,j)}{D^2 N(N-1)}. \tag{3}$$

We denote the product $(R-1)L$ as ΔE^*. Thus ΔE^* is a natural measurement to characterize the impact of a link on the *efficiency*: if ΔE^* has a negative correlation with *range*, then short-range attacks are more destructive; and if ΔE^* has a positive correlation with *range*, then long-range attacks are more destructive.

For the WS model, most links are local connections, having a short *range* and a small *load*, while a few links are rewired links, having a long *range* and a large *load*. There is a clear positive correlation between *load* and *range*, thus ΔE^* has a positive correlation with *range*. As a result, long-range attacks are more destructive for the WS model.

For many RSF networks, short-*range* links tend to link together nodes with high degree, and these links are expected to be passed through by a large number of shortest paths. Thus high *load* is associated mainly with short-range links [3]. The relation between L and R, and the according relation between ΔE^* and R are plotted in figure 1. From figure 1, we can see ΔE^* has a negative correlation with R, which indicates that for RSF networks, short-range attacks are more destructive.

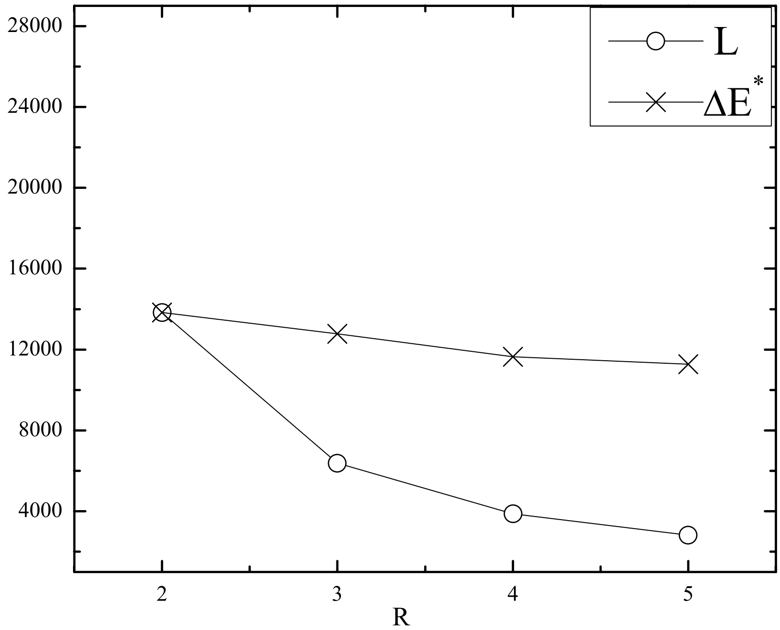

Figure 1. Averaged *load* and ΔE^* as a function of *range* for the RSF networks with $\lambda = 3$, $N = 5000$, minimal degree $k_0 = 3$ and maximal degree $k_m = 500$. The data of *range–load* is from [3], and the averaged ΔE^* is computed by us based on their data directly.

Since many real networks, e.g. the Internet and power grids etc, exist in two- or three-dimensional physical spaces, it is natural to study the geographical complex networks, and this has attracted much attention recently [8]–[19]. In this paper, we consider how the geography influences the efficiency of a network. Here, a weighted lattice embedded scale-free (WLESF) network model is considered [17]. The WLESF network is generated as follows: (i) a lattice with periodic boundary conditions of size $L \times L$ is assumed, upon which networks will be embedded; (ii) for each node an integer k is assigned as the largest degree it could have, keeping in mind that the distribution of k is a power law function: $P(k) \sim k^{-\lambda}$, $k_0 \leqslant k \leqslant k_m$; (iii) a node is randomly selected (say, i, with degree k_i) from the lattice, then according to a Gaussian weight function $f_i(r) = D\mathrm{e}^{-(r/A\sqrt{k_i})^2}$, it selects other nodes (say, j) and establishes a connection between them if j's quota is not filled yet and there exists no previous connection between i and j, until its degree quota k_i is filled or until it has tried $1/f_i(3R_i)$ times; (iv) the process is repeated throughout all the nodes on the lattice. The normalization constant D, defined by $\int_1^\infty \mathrm{d}r 2\pi f_i(r) = 1$, is $(\pi A^2 k_i)^{-1}\mathrm{e}^{1/A^2 k_i}$, and $R_i = A \times \sqrt{k_i}$, serves as the characteristic radius of the region that node i can almost freely connect. In this model, the cutoff parameter A controls the strength of geographical constrains. We plot the distribution of *range* for WLESF networks with $A = 1$, 2 and 20 in figure 2, and the clustering coefficient C as a function of A in the inset of figure 2. From this figure, we can see that for small A, most links are localized, and the networks have a large clustering coefficient and a small average *range*, similar to the WS model other than the degree distribution. For large A, the networks have a small clustering coefficient and a large average *range*, like a RSF network [20]. Figure 2 shows the case of $\lambda = 3.5$, other cases have the same behavior.

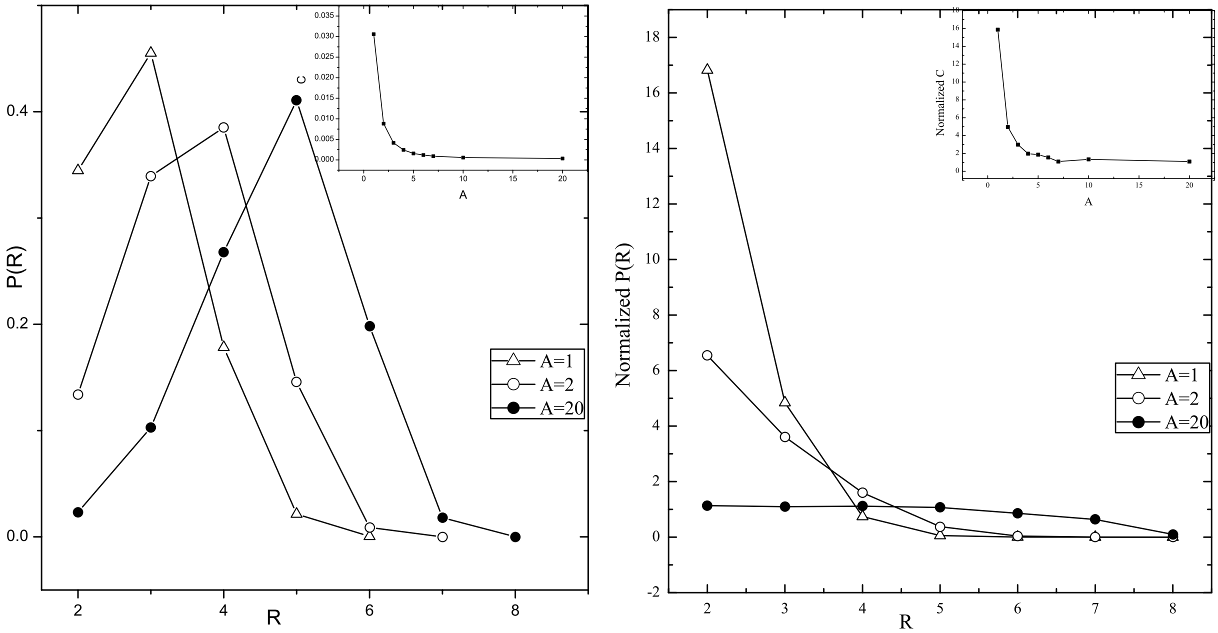

Figure 2. Distribution of *range* $P(R)$ for WLESF networks with $\lambda = 3.5$, $k_0 = 6$, $k_m = 500$, $N = 10\,000$. Inset: clustering coefficient C as a function of A for the same network. The quantities in the right figure are normalized by that of the corresponding RSF network.

The aim of this paper is to investigate explicitly the impact of the geographical constraints by analyzing the range-based attacks on scale-free networks. The studying of attacks on complex networks has been an issue for a long time [21]–[24]. Attack here is defined as the deliberate removal of a subset of selected links. We do both short- and long-range attacks on links in WLESF networks for different values of λ and A. For short-range attacks, links with shorter *ranges* are removed first; for long-range attacks, links with longer *ranges* are removed first. In both cases, the choice among links with the same *range* is made at random. We measure the *efficiency* of the network as links are successively removed, and plot the *efficiency* (normalized by its initial value) as a function of the fraction of removed links in figure 3. For a wide range of λ, as shown in figures 3(b)–(e) for $\lambda = 2.95$, 3.0, 3.5 and 5.0, respectively, short-range attacks are more destructive when the geographical constraints are weak ($A = 20$ for illustration). As the geographical constraints become stronger, long-range attacks become more vital ($A = 1$ for illustration). Here, we only show some typical values of A as illustrations. In fact, the roles of attacks on short- and long-range links switch continuously as the geographical constraints vary continuously. And there is a critical value A_c for which short- and long-range attacks have identical destructive effects on the network. A_c for networks with different λ are listed in table 1. When λ is sufficiently small or large, long-range attacks are always more destructive, and there is no A_c in our simulations.

The above results can be understood in the following. When the geographical constraints are strong, the system is composed of a number of clusters, where nodes within each cluster are densely connected, but the linkage among the clusters is sparse [17]. The long-range links are those linking different clusters while the short-range links are within individual clusters, thus long-range links play a key role for the efficiency of the system. When the geographical

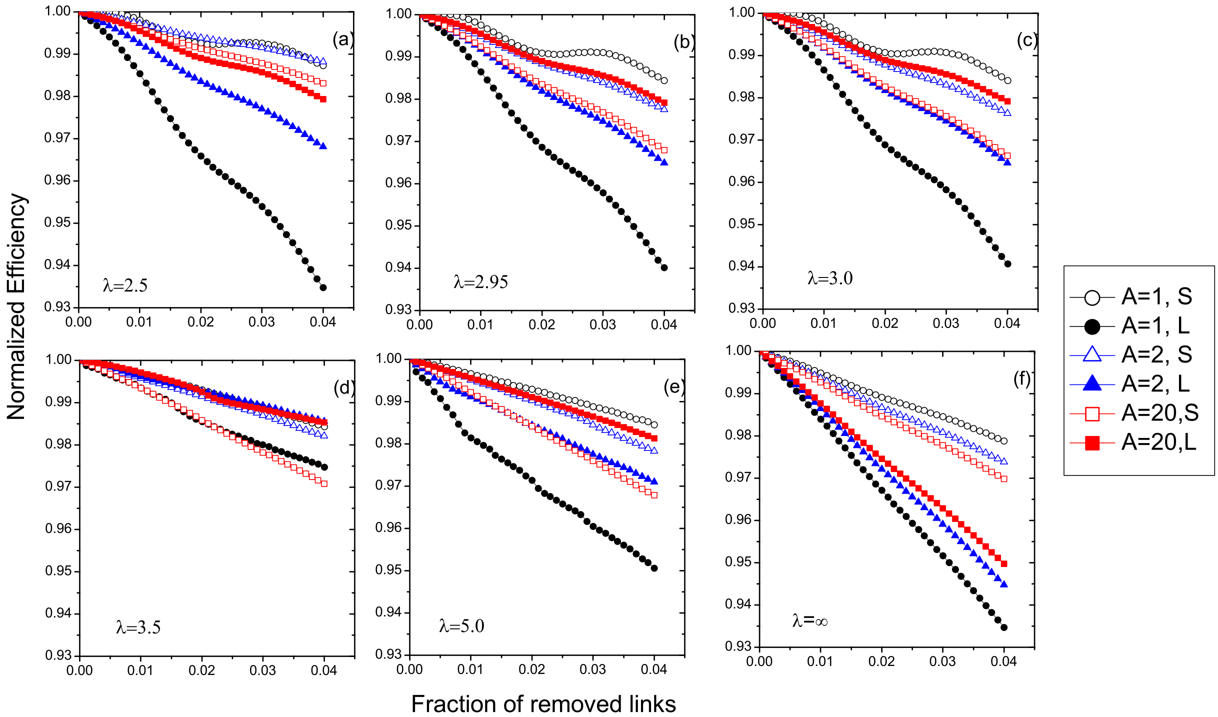

Figure 3. Normalized *efficiency* for short- (S) and long-range (L) attacks as a function of the fraction of removed edges in WLESF networks with $k_0 = 6$, $k_m = 500$, $N = 10\,000$ and different values of λ: (a) 2.5, (b) 2.95, (c) 3.0, (d) 3.5, (e) 5.0 and (f) ∞. Each data is averaged over 100 realizations.

Table 1. Critical value of geographical constraints A_c for WLESF networks with different values of λ.

λ	2.5	2.95	3.0	3.5	5.0	∞
A_c	None	2.5	2.56	1.95	2.8	None

constraints are weak, the system is approximately a RSF network. Short-range links tend to link together highly connected nodes, while long-range links tend to connect nodes with very few links. Thus, short-range links are more important than the long-range links for the efficiency of the network. A quantitive confirmation can been seen from the relation between ΔE^* and R. We plot the $L - R$ and $\Delta E^* - R$ relations in figure 4. From this figure, we can see that for WLESF networks, when geographical constraints are weak, both L and ΔE^* have a negative correlation with R, meaning that short-*range* links is associated mainly with high *load*, and contribute to the efficiency of the network more than long-*range* links associated mainly with low *load*. When geographical constraints increase, the correlation between ΔE^* and R becomes positive, meaning long-*range* links contribute more to the efficiency of the network, similar to the WS model.

Two exceptions occur for networks with sufficiently small or large values of λ, for which long-range attacks are always more destructive when changing the geographical constraints. In figures 3(a) and (f), we show the results for $\lambda = 2.5$ and $\lambda = \infty$, respectively. The two exceptions

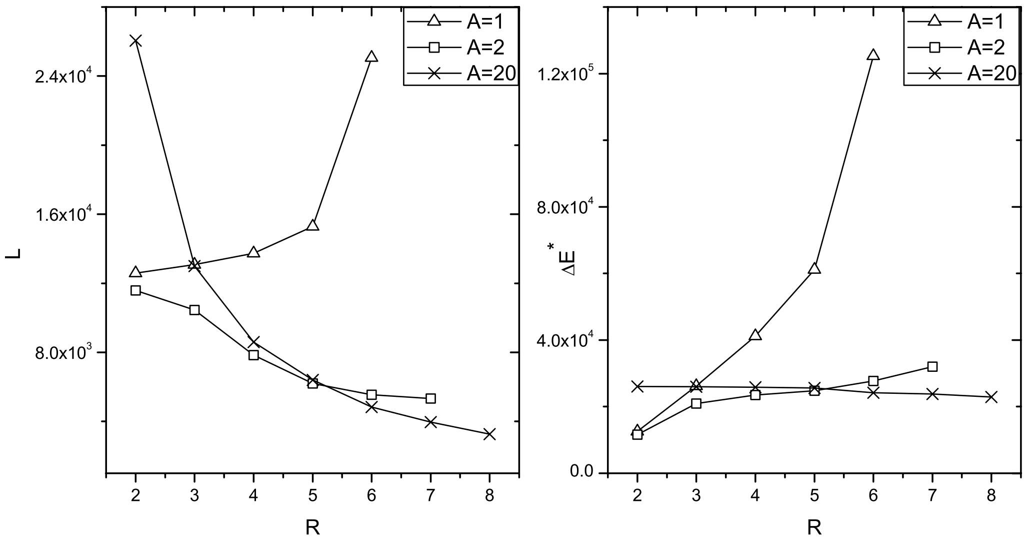

Figure 4. Averaged *load* (left panel) and average ΔE^* (right panel) as a function of *range* for WLESF networks with $A = 1$, 2 and 20. All the parameters are the same as in figure 2.

in the case of RSF networks are reported by Motter *et al* [3]. For networks with small values of λ, there is a densely connected subnetwork of nodes with large connectivity; thus there are so many redundant short-range connections that the removal of one will not increase the average shortest path to a great extent. For networks with large values of λ, switching of the roles of short- and long-range attacks is caused by the homogenization of the network, where all the nodes have approximately the same connectivity, thus links with higher load are precisely those between distant nodes, i.e. those with larger range [3]. For WLESF networks with sufficiently small or large value of λ, our results suggest that geographical constraints make long-range links even more important than those in RSF networks.

In summary, we discuss range-based attacks on links in geographically constrained networks, and show that when geographical constraints are weak, short-range attacks are more destructive, and when geographical constraints increase, long-range links become more and more vital. Our results demonstrate that geography has a significant impact on the efficiency and robustness of complex networks. Based on the analysis in [4], we introduce a quantity ΔE^* in this paper, which give a reasonable measure of the impact of links on the *efficiency* of the network, thus one may design a ΔE^*-based attacking strategy, which should be more effective for both WS models and RSF networks. Our results may be instructive for both construction and destruction of complex networks.

Acknowledgments

This work is supported by the National Hi-Tech Research and Development Program with grant no. 2006AA09A102-08 and the National Basic Research Program of China with grant no. 2007CB209603. L Y thanks the 100 Person Project of the Chinese Academy of Sciences, and the China National Natural Science Foundation with grant no. 10775157 for their support.

References

[1] Watts D J and Strogatz S H 1998 *Nature* **393** 440
[2] Watts D J 1999 *Small World* (Princeton, NJ: Princeton University Press)
[3] Motter A E, Nishikawa T and Lai Y 2002 *Phys. Rev.* E **66** 065103
[4] Gong B, Liu J, Huang L, Yang K and Yang L 2007 *Preprint* 0708.0150
[5] Latora V and Marchiori M 2001 *Phys. Rev. Lett.* **87** 198701
[6] Newman M E J 2001 *Phys. Rev.* E **64** 016132
[7] Goh K I, Kahng B and Kim D 2001 *Phys. Rev. Lett.* **87** 278701
[8] Rozenfeld A F, Cohen R, ben-Avraham D and Havlin S 2002 *Phys. Rev. Lett.* **89** 218701
[9] ben-Avraham D, Rozenfeld A F, Cohen F and Havlin S 2003 *Physica* A **330** 107
[10] Warren C P, Sander L M and Sokolov I M 2002 *Phys. Rev.* E **66** 056105
[11] Dorogovtsev S N and Mendes J F F 2000 *Europhys. Lett.* **50** 1
[12] Dall J and Christensen M 2002 *Phys. Rev.* E **66** 016121
[13] Nemeth G and Vattay G 2003 *Phys. Rev.* E **67** 036110
[14] Herrmann C, Barthélemy M and Provero P 2003 *Phys. Rev.* E **68** 026128
[15] Kaiser M and Hilgetag C C 2004 *Phys. Rev.* E **69** 036103
[16] i Cancho R F 2004 *Phys. Rev.* E **70** 056135
[17] Yang K, Huang L and Yang L 2004 *Phys. Rev.* E **70** 015102
[18] Huang L, Yang L and Yang K 2005 *Europhys. Lett.* **72** 144
[19] Huang L, Yang L and Yang K 2006 *Phys. Rev.* E **73** 036102
[20] Newman M E J, Strogatz S H and Watts D J 2001 *Phys. Rev.* E **64** 026118
[21] Albert R, Jeong H and Barabasi A L 2000 *Nature* **406** 378
[22] Cohen R, Erez K, ben-Avraham D and Havlin S 2000 *Phys. Rev. Lett.* **85** 4626
[23] Callaway D S, Newman M E J, Strogatz S H and Watts D J 2000 *Phys. Rev. Lett.* **85** 5468
[24] Cohen R, Erez K, ben-Avraham D and Havlin S 2001 *Phys. Rev. Lett.* **86** 3682