

# 地震数据关系网络的空间尺度

何璇<sup>1</sup>, 王卢阳<sup>2</sup>, 赵海<sup>2</sup>, 刘晓<sup>2</sup>

(1. 东北大学 中荷生物医学与信息工程学院, 辽宁 沈阳 110169; 2. 东北大学 计算机学院, 辽宁 沈阳 110169)

**摘 要:** 通过分析美国加州地区地震数据关系网络的拓扑结构特征,研究了不同空间尺度下网络拓扑结构基本度量参数的变化规律. 结果表明:随着空间尺度的增加,网络规模迅速变小,并最终达到稳定;网络的平均聚集系数和平均最短路径长度随空间尺度的增加呈波动性下降. 度分布频度研究结果表明,网络的空间尺度在一个有效范围内时,网络呈现出较好的幂律分布,该范围受事件数量影响.

**关 键 词:** 复杂网络;地震数据关系网络;空间尺度;拓扑结构;幂律分布

**中图分类号:** TP 393.17      **文献标志码:** A      **文章编号:** 1005-3026(2018)10-1385-05

## Space Scales of Earthquake Data Relation Networks

HE Xuan<sup>1</sup>, WANG Lu-yang<sup>2</sup>, ZHAO Hai<sup>2</sup>, LIU Xiao<sup>2</sup>

(1. School of Sino-Dutch Biomedical & Information Engineering, Northeastern University, Shenyang 110169, China; 2. School of Computer Science & Engineering, Northeastern University, Shenyang 110169, China.

Corresponding author: HE Xuan, E-mail: hexuan@bmie.neu.edu.cn)

**Abstract:** By analyzing the topological structural characteristics of the earthquake data relation network in the California region of the United States, the changing rules of the basic metric parameters of the network topological structure under different space scales were studied. The results showed that with the increase of the space scales, the network scale rapidly becomes smaller and eventually reaches stability. The average clustering coefficient and the average shortest path length of the network decrease with fluctuations as space scales increase. The results of degree distribution frequency showed that when the space scales of the network are within a valid range, the network presents a better power-law distribution, which is affected by the number of events.

**Key words:** complex network; earthquake data relation network; space scales; topological structure; power-law distribution

网络的拓扑结构分析是复杂网络研究中非常基础且重要的部分. 在复杂网络上的合作、博弈、同步与动力学行为分析等方面的研究成果很大程度上受到了网络拓扑结构的影响,不同拓扑结构的网络动力学行为表现出非常明显的差异.

在由地震数据关系生成的网络拓扑中,节点通常代表所研究的地理区域内被划分成大小相等的区域单元,边通常代表区域单元之间的影响关系. 2003年,日本 Abe 团队针对南加州 1984~2001 年的地震数据进行了分析,发现了地震事件之间存在复杂的相互影响关系. 2004年, Abe 团队首次将复杂网络的方法引入到地震研究中,并

以南加州地区的地震数据为研究对象,提出一种基于时间序列的地震数据关系网络构造方法<sup>[1]</sup>. 从 2004 年至今, Abe 团队坚持不懈地研究地震数据关系网络的复杂网络特性,并发表了一系列研究成果<sup>[1-4]</sup>. 证明了加州地区的地震数据关系网络具有无标度<sup>[1]</sup>和小世界<sup>[2]</sup>特征,发现了相继发生的地震事件时间间隔符合幂率分布<sup>[3]</sup>,以及地震数据关系网络具有层次结构和同配性<sup>[4]</sup>,并从侧面证明了小地震在地震数据关系网络拓扑结构中的重要性.

除 Abe 团队之外,研究人员分别在地震数据关系网络建模方法、网络拓扑特征(层次性、聚集

性、周期性)等方面展开研究<sup>[5-11]</sup>. 其中,空间尺度的讨论是其中一个非常重要的方面. 文献[10-11]指出尺度选择对于网络研究结果的影响,并进行了相应的讨论,发现网络尺度的改变对网络的规模及网络的基本拓扑特征参量均会产生影响,并确定了空间尺度的下限,这对地震网络的进一步研究起到了非常重要的作用.

在前期的研究中,本课题组提出了一种基于时空影响域的地震网络构造方法<sup>[12]</sup>. 该方法同时考虑了地震事件在时间和空间上的影响关系,并且所得网络的度分布符合幂律分布. 以上研究中,课题组将所研究的空间尺度设为 5~20 km 之间. 课题组在后续的研究中,发现尺度特别大或者特别小对网络特征的研究结果都会产生影响. 因此本文特别针对时空影响域网络,研究在不同空间尺度下的网络拓扑结构及其特征参数的变化情况,最终寻找地震数据关系网络的有效空间尺度.

## 1 空间尺度

一般来说,尺度是研究客体或过程的空间维和时间维,可用分辨率和范围来描述,它标志着对所研究对象细节了解的水平,空间尺度通常是指最小的可辨识单元所代表的特征长度、面积或体积,地理空间数据具有尺度依赖性. 比如,人们熟知的大陆漂移说,之所以被发现,就是因为著名地球物理学家阿尔弗雷德·魏格纳观察了一张大尺度的完整地图,而不是观察小尺度下一个微小的细节. 可以想象,如果在小尺度空间中,这个轰动世界的学说是不可能被发现的. 因此,在很多研究领域,把研究实体用尺度不同的单元划分后,其研究结果会发生变化,空间对象的复杂程度也会发生变化.

图 1 是在不同空间尺度划分下的网络拓扑结构. 其中, $L$  为最小单元的边长,在同一单元内的节点均视为同一节点. 随着  $L$  的变化,可见拓扑结构随之改变. 发现  $L=1$  的拓扑结构和原始拓扑结构比较近似;然而,当  $L=2$  时,节点数减少一半,观察角度更宏观. 极端情况下,当  $L=4$  时,原图由带自连边的一个节点所代替,完全掩盖了拓扑结构的所有信息.

因此,研究人员希望可以获得地理单元的最佳空间尺度. 尺度过大,则会掩盖大量有价值的信息;尺度过小,无法获得地理单元所代表的特征. 比如,对于地震数据关系网络,从精确计算的角度而言,将一个具有准确位置的地震事件作为地震

网络中的一个节点是可行的. 然而,如果想理解地震事件之间存在何种关系,就必须研究地震事件所在地理单元的特点. 换言之,单一发生的地震事件无法代表所在地理单元内发生的所有地震事件.

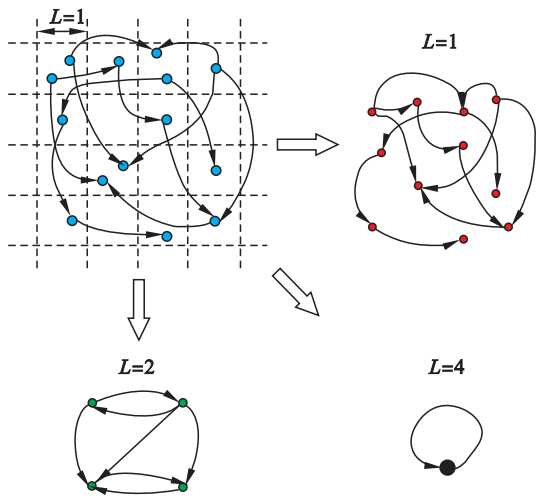


图 1 不同空间尺度下的网络拓扑结构  
Fig. 1 Network topological structure at different scales

## 2 数据来源与相关定义

### 2.1 数据来源

世界范围内有许多大大小小的地震数据中心,其中,美国南加州地震中心(Southern California Seismic Network, SCSN)提供从 1932 年至今超过 80 年的加州地区地震数据,拥有超过 50 万条公开地震记录. 随着数据采集和测量方法的不断改进,SCSN 在不同的地方设有超过 400 个探测站. 尤其可以探测到那些较小的地震(1.0 级以下),为地震研究提供了非常重要的支持. 采用该中心提供的数据进行空间尺度分析,选择 1992~2001 年共计 10 年的地震数据,地理空间范围为:北纬  $32^{\circ}\sim 36^{\circ}$ ,西经  $114^{\circ}\sim 122^{\circ}$ ,共计发生 222 110 次地震. 图 2 为加州地区使用时空影响域方法生成的地震数据关系网络拓扑结构图.

### 2.2 相关定义

本文将根据以下网络拓扑特征参数对不同空间尺度下的网络拓扑结构进行研究. 具体定义如下:

**定义 1** 度与平均度:网络中一个节点的连接数被定义为该节点的度值,以  $k$  表示. 通常情况下,节点的度值越大,它对整个网络的影响也就越大. 网络中所有节点度值的平均值称为平均度,用  $\langle k \rangle$  表示.

**定义 2** 度分布频度:网络中节点的度分布

用  $P(k)$  来表示,表示任意一个节点度值为  $k$  的概率.在规则网络中,节点具有大致相同的度值,度分布为 Delta 分布.随机网络的度分布则近似于 Poisson 分布.近年来的研究表明,许多实际网络的度分布为幂律分布.幂律分布描述了复杂网络中节点度值分布的异质性.对于幂律分布有  $P(k) \sim k^{-\gamma}$ ,  $\gamma$  为幂指数.

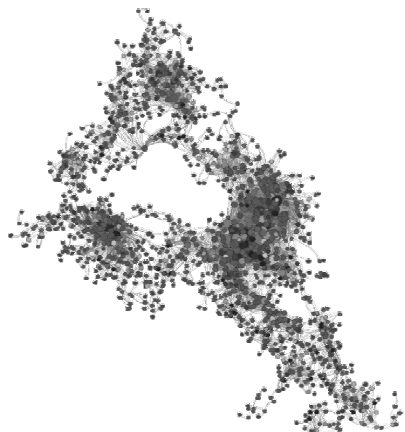


图 2 基于时空影响域的加州地区地震网络拓扑结构图  
Fig. 2 Earthquake network topological structure of California based on space-time influence domain

**定义 3** 平均最短路径长度:网络中任意两个节点  $v_i$  和  $v_j$  的最短路径为连接这两个节点边数最少的路径.距离  $d_{i,j}$  被定义为连接这两个节点的最短路径的边数.网络的平均最短路径长度则被定义为任意两个节点之间距离的平均值,即

$$D = \frac{2}{n(n+1)} \sum_{i=1, j>i}^n d_{i,j}. \quad (1)$$

平均最短路径长度通常代表了网络的传输效率.其值越小,代表网络的传输效率越高.

**定义 4** 平均聚集系数:一个节点的聚集系数  $C_i$  是它的相邻节点之间的连接数与它们所有可能存在连接的数量的比值,即

$$C_i = \frac{2x}{n(n-1)}. \quad (2)$$

式中:  $x$  是这个节点与相邻节点之间的连接数;  $n$  为这个节点相邻节点的个数.网络的平均聚集系数即为网络中所有节点聚集系数的平均值,即  $C = \frac{1}{n} \sum_{i=1}^n C_i$ .聚集系数描述了网络中节点的聚集程度,通常与平均最短路径长度一起衡量实际网络是否具有小世界特征.

### 3 多尺度网络特征分析

本文分别按照不同的  $L(4 \sim 50 \text{ km})$  划分所研

究的地理区域,并按照时空影响域的方法生成网络,得到 47 个不同的关系网络.计算不同关系网络的拓扑特征参数,并进行比较.

节点数和边数随单元长度的变化结果如图 3 所示,在  $L$  增大的过程中,网络的节点数和边数均呈下降趋势.该网络在尺度大于 20 km 后,规模几乎不变.

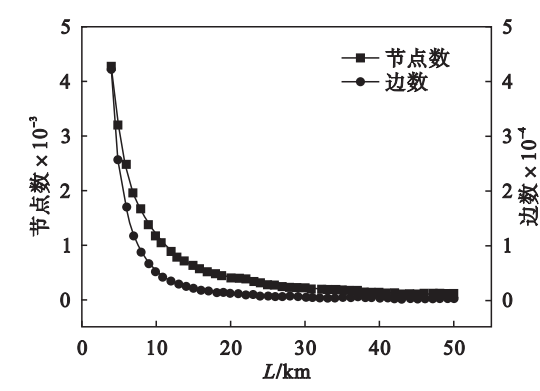


图 3 节点数和边数随单元长度  $L$  的变化  
Fig. 3 Change of number of nodes and edges with cell size  $L$

网络的平均度、平均聚集系数及平均最短路径长度随单元长度的变化如图 4 所示.结果表明,随着  $L$  的增大,网络的平均度值(图 4a)逐渐降低,在  $L$  值大于 20 km 后,平均度值几乎保持不变,说明网络的密度达到了一个相对稳定的程度.随着  $L$  的增大,网络的平均聚集系数(图 4b)呈现波动性的下降趋势.说明网络的聚集程度在尺度逐渐增大的过程中逐渐变小,然而这种变化与尺度的变化并不具备严格的线性相关.随着  $L$  的增大,网络的平均最短路径长度(图 4c)呈现波动性的下降趋势,但总体上平均最短路径长度在 5 ~ 10 之间,结合平均聚集系数的相关结果和判断小世界的方法,说明随着  $L$  的变化,网络的小世界特征并没有消失,这从侧面印证了地震活动的长程相关特征.

需要注意的是,与本文不同,在文献[10-11]得到的研究结果中,网络的平均聚集系数和平均路径长度随单元长度的增加逐渐减小.研究发现,根本原因在于建网方法的区别.如前文所述,本文建网方法仅考虑地震事件之间的直接联系.当  $L$  增大时,距离较近的节点会合并成同一个节点,它们之间的连边被吞并,从网络的形式上来看,最终变成了自连边.而自连边会在网络预处理步骤中被删除.然而,Abe 的建网方法可以产生距离非常远的节点之间的连边.当  $L$  增大时,距离较近节点合并成一个节点,然而距离非常远的

连边却仍然存在,因此网络变得更聚集.

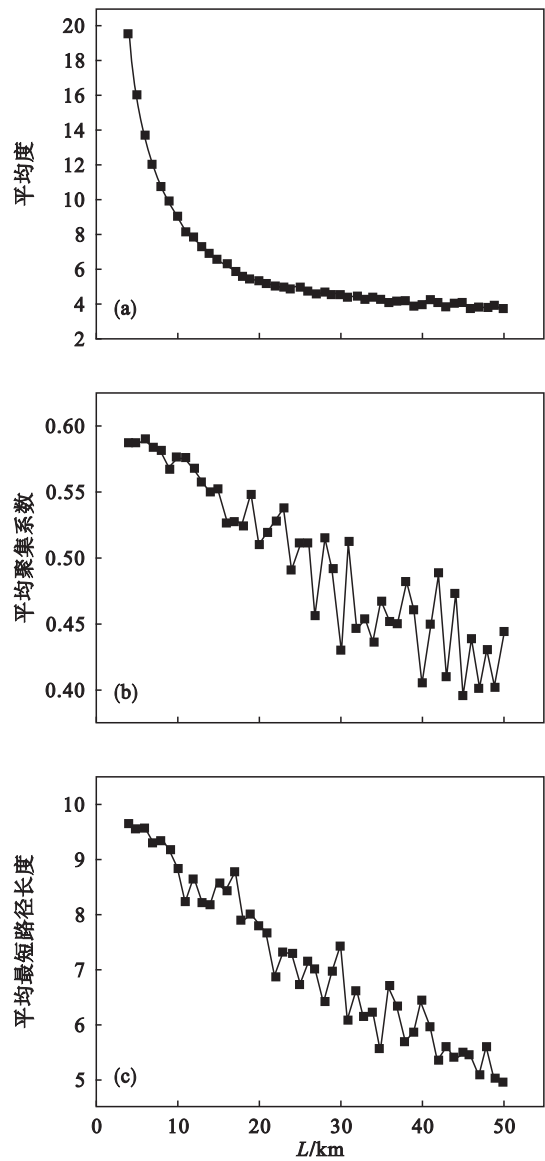


图 4 网络的不同拓扑特征量随单元长度的变化  
Fig. 4 Change of topological features of network with cell size

按照文献[13]的方法计算了幂率分布指数  $\gamma$  和参数  $p$ . 文献中认为当  $p$  的值大于 0.1 时,网络表现出较好的幂律特征. 不同空间尺度下的网络度分布幂指数  $\gamma$  和  $p$  如图 5 所示,随着  $L$  的增大,网络的幂指数  $\gamma$  并没有呈现明显的变化规律. 然而,  $p$  值却在  $L \in [9, 33]$  的范围内始终大于 0.1, 说明当网络的单元长度在这个范围内时表现出较好的幂率分布特征. 与 Abe 等的结果不同, 本文不仅确定了  $L$  取值的下限, 也确定了  $L$  取值的上限. 因此, 在本文的网络生成方法中,  $L$  不能无限增大.

$L$  的有效取值范围受事件数量的影响, 不同事件数下  $L$  的有效取值范围如图 6 所示. 当选取事件总数分别为 50 189 和 371 516 时, 得到  $L$  的

有效取值范围分别为  $[5, 26]$  和  $[10, 35]$ . 因此, 按照上述方法, 对于不同的事件数需要确定  $L$  的有效取值范围, 确保在一个有效范围内研究网络的拓扑特征量, 才能排除  $L$  的取值对于结果的影响.

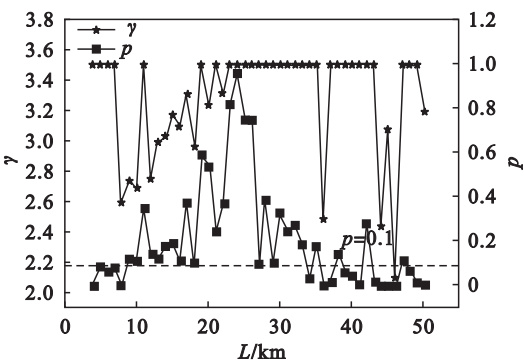


图 5 不同空间尺度下的网络度分布幂指数  $\gamma$  和  $p$   
Fig. 5 Network degree distribution exponent  $\gamma$  and  $p$  for different space scales

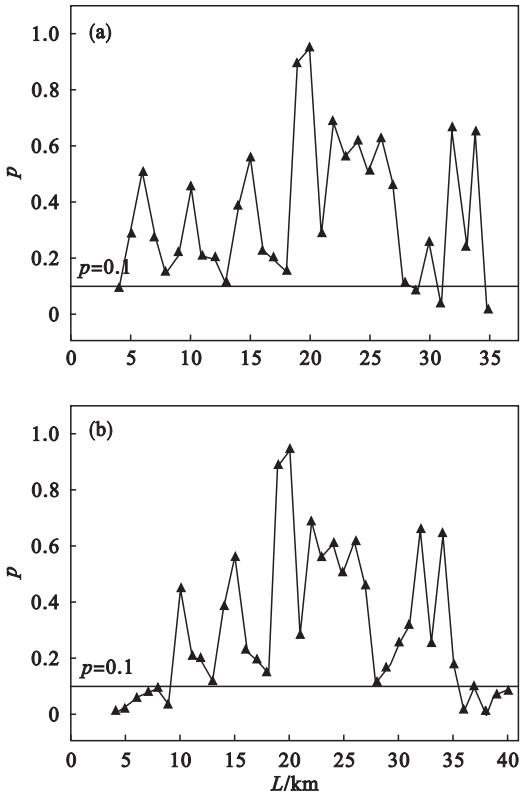


图 6 不同事件数下  $L$  的有效取值范围  
Fig. 6 Valid ranges for different numbers of events  
(a)—50 189; (b)—371 516.

## 4 结 论

1) 针对基于时空影响域的地震数据关系网络, 研究空间尺度的改变对于网络拓扑结构的影响, 发现空间尺度的改变会影响网络拓扑特征参数.



2) 发现了网络拓扑特征参数随尺度变化所呈现的规律:随着尺度的增大,网络的规模开始迅速减小,而后几乎保持不变;网络的平均聚集系数和平均最短路径长度呈波动性下降;网络的幂指数  $\gamma$  并没有呈现明显的变化规律. 说明空间尺度的改变,会影响网络的规模、聚集程度和网络动力学效率.

3) 提出通过幂律分布的  $p$  值确定合理的空间尺度范围的方法,与以往方法相比,既确定了空间尺度的下限,又确定了空间尺度的上限.

参考文献：

[ 1 ] Abe S, Suzuki N. Scale-free network of earthquakes [ J ]. *Europhysics Letters*, 2004, 65 ( 65 ) : 581 – 586.

[ 2 ] Abe S, Suzuki N. Small-world structure of earthquake network [ J ]. *Physica A : Statistical Mechanics & Its Applications*, 2004, 337 ( 1/2 ) : 357 – 362.

[ 3 ] Abe S, Suzuki N. Statistical similarities between internet quakes and earthquakes [ J ]. *Physica D : Nonlinear Phenomena*, 2004, 193 ( 1 ) : 310 – 314.

[ 4 ] Abe S, Suzuki N. Complex earthquake networks: hierarchical organization and assortative mixing [ J ]. *Physical Review E : Statistical Nonlinear & Soft Matter Physics*, 2006, 74 ( 2 ) : 026113.

[ 5 ] He X, Zhao H, Cai W, et al. Analyzing the structure of earthquake network by  $k$ -core decomposition [ J ]. *Physica A: Statistical Mechanics & Its Applications*, 2015, 421 ( 1 ) : 34 – 43.

[ 6 ] Zhang Y, Zhao H, He X, et al. Bayesian prediction of earthquake network based on space-time influence domain [ J ]. *Physica A: Statistical Mechanics & Its Applications*, 2016, 445 ( 1 ) : 138 – 149.

[ 7 ] Rezaei S, Darooneh A H, Lotfi N, et al. The earthquakes

network: retrieving the empirical seismological laws [ J ]. *Physica A: Statistical Mechanics & Its Applications*, 2017, 471 ( 1 ) : 80 – 87.

[ 8 ] Lin M, Fan X X, Wang G, et al. Network structure entropy and its dynamical evolution for recurrence networks from earthquake magnitude time series [ J ]. *The European Physical Journal B*, 2016, 89 ( 5 ) : 1 – 7.

[ 9 ] 徐久强, 宋佳, 何璇, 等. 基于 OFC 的仿真地震序列网络化特征对比分析 [ J ]. *东北大学学报 ( 自然科学版 )*, 2017, 38 ( 2 ) : 205 – 208.

( Xu Jiu-qiang, Song Jia, He Xuan, et al. Comparison analysis of network characteristics of OFC-based simulated earthquake sequences [ J ]. *Journal of Northeastern University ( Natural Science )*, 2017, 38 ( 2 ) : 205 – 208. )

[ 10 ] Abe S, Suzuki N. Determination of the scale of coarse graining in earthquake network [ J ]. *Europhysics Letters*, 2009, 87 ( 4 ) : 48008 – 48012.

[ 11 ] Lotfi N, Darooneh A H. The earthquakes network: the role of cell size [ J ]. *The European Physical Journal B*, 2012, 85 ( 1 ) : 1 – 4.

[ 12 ] 何璇, 赵海, 蔡巍, 等. 基于时空影响域的地震网络构造方法 [ J ]. *东北大学学报 ( 自然科学版 )*, 2014, 35 ( 10 ) : 1395 – 1399.

( He Xuan, Zhao Hai, Cai Wei, et al. Construction method based on time-space influence domain for earthquake network [ J ]. *Journal of Northeastern University ( Natural Science )*, 2014, 35 ( 10 ) : 1395 – 1399. )

[ 13 ] Clauset A, Shalizi C R, Newman M E J. Power-law distributions in empirical data [ J ]. *Society for Industrial and Applied Mathematics*, 2009, 51 ( 4 ) : 661 – 670.