

DOI:10.3876/j.issn.1000-1980.2014.02.015

基于复杂网络理论的河流网络建模与分析

吴学文,瞿永钢,李 玲

(河海大学计算机与信息学院,江苏 南京 210098)

摘要:应用复杂网络理论建立河流网络模型,并以海河水系为例,建立了 2 个网络模型:模型 I 和模型 II。模型 I 只包含自然节点;而模型 II 既包含自然节点,也包含人工节点。针对河流网络节点和边的自身特殊性,提出了河流网络节点重要性的权值指标和边重要性的权值指标,并研究和比较海河网络的节点及边的重要性;最后,对所建河网模型和提出的重要参数指标进行总结与展望。

关键词:复杂网络;河流网络建模;节点重要性;边重要性;海河水系

中图分类号:TV212.4 文献标志码:A 文章编号:1000-1980(2014)02-0177-06

River network modeling and analysis based on complex network theory

WU Xuewen, QU Yonggang, LI Ling

(College of Computer and Information Technology Engineering, Hohai University, Nanjing 210098, China)

Abstract: In this study, we developed a river network model based on the complex network theory. Taking the Haihe River system as an example, we constructed two river network models, Model I and Model II. Model I included natural vertices and Model II included both natural vertices and artificial vertices. Based on the properties of the vertices and edges in river networks, we propose the vertex weight index and edge weight index to assess the vertex importance and edge importance, respectively, and analyze and compare the vertex importance and edge importance of the Haihe River Network. Finally, we summarize the key parameters of the proposed river network model and discuss the future development of the model.

Key words: complex network; river network modeling; vertex importance; edge importance; Haihe River system

目前,复杂网络理论应用于各种学科领域。倪向萍等^[1]应用复杂网络理论研究了电力网络的脆弱度;赵月等^[2]将复杂网络应用于城市交通网络研究;Sporns^[3]用复杂网络理论研究人脑连接组并研究网络动态特性;田占伟等^[4]发现微博信息传播网络具有集群性、小世界、高度中心化等特征。

河流网络是一个远离平衡态的、自组织的、开放的网络^[5-6],是一个典型的复杂网络。复杂网络理论为研究河流网络提供了新思路和新方法。20 世纪 90 年代,学者开始应用复杂网络理论来研究河流网的形成及演化机理^[7]。1992—1993 年间,Rodriguez- I turbe 等^[8]、Rinaldo 等^[9]应用“哈密顿量最小化选择生成树”所得到的“最佳通道网”可以重现实际的河流网络形态。2004—2005 年间,沈丹等^[10]、许田^[11]、汪小帆等^[12]研究了长江流域地理网络的自组织临界现象,证明各级支流的总流域面积和各级支流的总长度对河流级别的依赖关系都符合无标度规律。2005 年,He 等^[13]证明了北京-廊坊-天津地区的天然河流网络服从指数为 2.48 的无标度分布。

笔者应用复杂网络理论来建立河流网络模型,并进一步研究节点及边的重要性。该研究可以为水利规划、洪水调度提供理论依据。

收稿日期:2013-09-23

基金项目:国家自然科学基金(51179047)

作者简介:吴学文(1962—),男,江西进贤人,副教授,博士,主要从事复杂系统与复杂网络、网络化信息获取与处理研究。E-mail: hhuwxw@hhu.edu.cn

1 基于复杂网络理论的河流网络建模

1.1 河流网络模型构建

河流网络可以抽象概化为一个有向加权网络模型(G),其基本要素主要包括节点集合(V)、边集合(E)、权重向量集合(W),即

$$G = \{V, E, W\}$$
$$V = \{v_1, v_2, \cdots, v_n\} \quad E = \{ \langle v_i, v_j \rangle \mid v_i, v_j \in V \}$$
$$W = \{w_{ij} \mid w_{ij} = (l_{ij}, q_{ij}(t)), v_i, v_j \in V\}$$

其中

网络节点定义为河流的流量、水位等水力特征明显改变的地方,包括河流源口、河流交汇点、河流出口、水电站、水库、水闸等;边是指河水在节点间流经的途径,即河道,包括自然河道和人工河道,边 $\langle v_i, v_j \rangle$ 表示一条从节点 v_i 到节点 v_j 的河道;权值是指各节点和各边的重要性。

1.2 应用实例

海河水系是中国华北地区最大的水系。根据《中华人民共和国水力资源复查成果》及其他相关资料^[14],将海河水系抽象概化为网络图,建立海河网络模型 I (图 1),其包括 344 个节点和 385 条边,只含自然节点和自然河道。为了研究工程节点、人工河道对河流网络的影响,建立了含有工程节点和人工运河的海河网络模型 II (图 2),其含有 565 个节点和 612 条边。

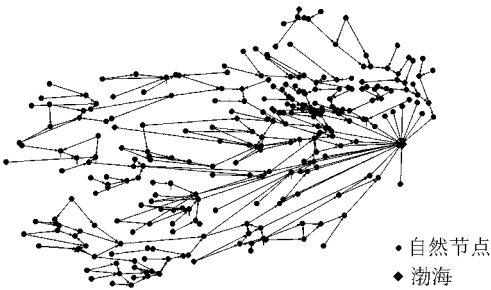


图1 海河网络模型 I
Fig. 1 Haihe River Network Model I

2 复杂河流网络的节点重要性

分析河网中节点及边的重要性并进行重点保护,有利于认识、预测和控制整个网络。目前,衡量节点重要性的指标主要有度指标和介数指标等。这里主要介绍更能反映网络状态的介数指标。

介数指标用来刻画网络中节点对于信息流动的影响力。节点 x 的介数指标定义为

$$C_B(x) = \frac{2 \sum_{j < k} g_{jk}(x)}{(n-1)(n-2)}$$

(2)

式中: $g_{jk}(x)$ —— g_{jk} 中经过节点 x 的最短路径数;
 g_{jk} ——节点 j 与节点 k 之间的最短路径数; $(n-1)(n-2)/2$ ——最大可能的节点介数(任意其他 2 个节点的最短路径都经过节点 x)。

2.1 河流网络节点重要性权值指标

由于河流网络节点具有其特有的性质,因此,本节提出了河流网络节点重要性权值指标。
河流网络中有自然节点和工程节点。不同类型的节点赋予不同的权值。所有的自然节点具有相同的类型权值 t_1 ,所有的工程节点具有相同的类型权值 t_2 。归一化的节点类型权值定义为

$$T_w(x) = \frac{T(x)}{T} = \frac{T(x)}{t_1 + t_2} = \begin{cases} \frac{t_1}{t_1 + t_2} & x \text{ 为自然节点} \\ \frac{t_2}{t_1 + t_2} & x \text{ 为工程节点} \end{cases}$$

(3)

式中: T ——自然节点和工程节点的权值之和, $T=t_1+t_2$ 。

节点的径流量反映节点在河流网络中的重要性。节点 x 的流量权值为



图2 海河网络模型 II
Fig. 2 Haihe River Network Model II

$$Q_w(x)=\frac{Q(x)}{Q_{\max}}$$

(4)

式中: $Q(x)$ ——节点 x 的多年平均径流量; $Q_{\max}$ ——河流网络中所有节点的多年平均径流量最大值。

河流网络中,2 个节点之间的连线对应于实体中的河道,连线长度对应于河道长度。连线长度体现了节点之间的紧密程度,连线越长,河道就越长,节点之间就越疏远。因此,将该节点到其余节点的最短河道长度的倒数之和定义为节点的距离权值。当节点之间有多条路径时,选择河道长度最短的路径来计算距离权值。节点的距离权值可表示为

$$D_w(x)=\sum_{y\neq x}\frac{1}{D_{xy}}$$

(5)

$$D_{xy}=\min_{xy}d_{xy}$$

(6)

式中: D_{xy} ——节点 x 到节点 y 的最短河道长度; d_{xy} ——节点 x 到节点 y 的一条路径长度。对于无向网络,有 $d_{xy}=d_{yx}$, $D_{xy}=D_{yx}$ 。

河流加权网络的节点重要性指标定义为类型权值、流量权值和距离权值的乘积,即

$$C_v(x)=T_w(x)Q_w(x)D_w(x)$$

(7)

2.2 海河网络中节点重要性分析

2.2.1 海河网络构建

忽略较小的节点和边后,建立了含有 78 个节点和 78 条边的简化海河网络模型,如图 3 所示。

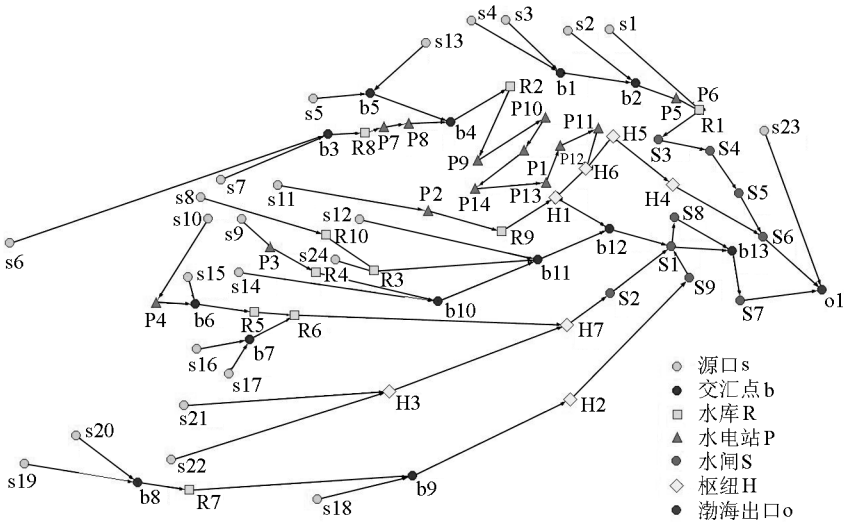


图 3 简化海河网络模型

Fig. 3 Simplified Haihe River Network Model

2.2.2 海河网络节点重要性分析

2.2.2.1 介数指标

计算海河网络节点重要性的介数指标,将其按区间段分为 7 个等级,如表 1 所示。

表 1 海河网络节点重要性的介数指标分布情况

Table 1 Distribution of betweenness index of vertex importance in Haihe River Network			
等级	介数指标	实际节点	节点数/个
1	0.0519	独流减河进洪闸	1
2	0.0410~0.0414	向阳口一级水电站、官厅水库、向阳口二级水电站	3
3	0.0328~0.0400	下马岭水电站、交汇点(b4)、安家庄一级水电站、安家庄二级水电站、下苇甸水电站	5
4	0.0205~0.0291	陈各庄水电站、交汇点(b13)、卢沟桥枢纽、锅底分洪闸、象光洞水电站等	9
5	0.0102~0.0195	筐儿港枢纽、黄壁庄水库、海河防潮闸、密云水库、交汇点(b9)、岗南水库、吴村闸等	18
6	0.0027~0.0096	交汇点(b5)、岳城水库、南里自洁节制闸、西大洋水库、王快水库、艾辛庄枢纽、紫荆关水电站等	17
7	0	所有河流源口和河流出口	25

2.2.2.2 权值指标

计算海河网络节点重要性的权值指标,将其按区间段分为8个等级,如表2所示。

表2 海河网络的节点权值指标分布情况

Table 2 Distribution of vertex weight index in Haihe River Network

等级	权值指标	实际节点	节点数/个
1	0.3844	独流减河进洪闸	1
2	0.2627	海河防潮闸	1
3	0.1216~0.1555	锅底分洪闸、下苇甸水电站、陈各庄水电站、西河闸、宁车沽防潮闸、渤海出口、交汇点(b13)	7
4	0.0525~0.0815	向阳口一级水电站、向阳口二级水电站、安家庄二级水电站、献县枢纽、安家庄一级水电站、密云水库等	10
5	0.0348~0.0495	南里自洁节制闸、屈家店枢纽、黄壁庄水库、交汇点(b12)、吴村闸、向阳闸、四女寺枢纽等	8
6	0.0101~0.0294	象光洞水电站、岗南水库、郭家里水电站、交汇点(b4)、新盖房枢纽、张坊水库等	21
7	0.0011~0.0097	交汇点(b10)、倒马关水库、紫荆关水电站、源口(s6)、源口(s1)、交汇点(b5)等	19
8	0.0003~0.0009	源口或交汇点	11

2.2.2.3 2种指标对比分析

在海河网络中,介数指标对节点的重要性划分较详细。介数指标能找到独流减河进洪闸、向阳口一级水电站、官厅水库等最重要的节点(等级1~2),这些节点的度不一定大,但对于整个网络却具有十分重要的作用,体现了介数指标的准确性。但介数指标认为河流源口和河流出口都具有相同的重要性,即都最不重要,因此仍有些片面。

节点权值指标同介数指标一样,能准确找到最重要的节点——独流减河进洪闸。串联于同一河道上的多个节点具有相近的权值指标,即具有相近的重要程度,这类似于介数指标。在权值指标中,源口间的权值指标不再都是零值,甚至大于交汇点的权值指标。因此,在对源口的分析上,权值指标优于介数指标,所以权值指标能更准确地找到河流网络的重要节点,但计算复杂度较高。

3 复杂河流网络的边重要性分析

分析边的重要性,找出那些相对重要的边(即河道),有利于认识、保护和控制重要的河道。

3.1 河流网络边重要性的权值指标

3.1.1 边介数指标

目前,研究边的重要性有多种方法,例如边介数指标^[15]。边介数指标通过计算边介数来描述边对网络的信息传递能力^[15]。边介数指标可以定义为

$$B_E(x)=\frac{2\sum_{j<k} \frac{g'_{jk}(x)}{g'_{jk}}}{(n-1)(n-2)}$$

(8)

式中: $g'_{jk}(x)$ —— g'_{jk} 中经过边 x 的最短路径数; g'_{jk} ——节点 j 与节点 k 之间的最短路径数; $(n-1)(n-2)/2$ ——最大可能的边介数(任意其他2个节点的最短路径都经过边 x)。

3.1.2 边权值指标

在河流网络中,连边代表着现实中的河道,而河道具有实际长度,长度权值越大则该边越具有较高的重要性。长度权值可以表示为

$$L_w(x)=\frac{L(x)}{L_{\max}}$$

(9)

式中: $L(x)$ ——边 x 的实际河道长度; $L_{\max}$ ——河流网络中的最长河道长度。

多年平均流量可用来衡量一条河流的水量传输能力。边的流量权值定义为

$$Q'_w(x)=\frac{Q'(x)}{Q'_{\max}}$$

(10)

式中: $Q'(x)$ ——边 x 对应河道的多年平均流量; $Q'_{\max}$ ——网络中河道多年平均流量的最大值。

边权值指标定义为边介数、边长度权值与边流量权值的乘积,可以表示为

$$C_E(x) = B_E(x)L_W(x)Q'_W(x)$$

(11)

3.2 海河网络边重要性分析

3.2.1 边介数指标

计算海河网络边重要性的介数指标,共有 45 个值,将其按区间段分为 7 个等级,如表 3 所示。

表 3 海河网络边重要性的介数指标分布情况

Table 3 Distribution of betweenness index of edge importance in Haihe River Network			
等级	介数指标	边	边数/条
1	4.510×10^{-2}	(R2, P9), (P9, P10)	2
2	$(4.100\sim4.450)\times10^{-2}$	(b4, R2), (P10, P1), (P1, P14), (P14, P13)	4
3	$(3.070\sim3.990)\times10^{-2}$	(S1, b13), (P13, P12), (P12, P11), (P11, H6)	4
4	$(2.050\sim2.880)\times10^{-2}$	(P6, b4), (b12, S1), (H7, S2), (S2, S1), (b11, b12)等	11
5	$(1.090\sim1.750)\times10^{-2}$	(b3, R8), (R1, S3), (P6, R1), (P5, P6), (b5, b4)等	19
6	$(1.000\sim9.600)\times10^{-3}$	(R3, b11), (b8, R7), (b1, b2), (S5, S6), (R4, b10)等	37
7	3.000×10^{-4}	(s23, o1)	1

3.2.2 边权值指标

经计算得到海河网络边重要性的权值指标,得到 78 个不同的值,即每条边的权值指标都不相等,将其按区间段分为 8 个等级,如表 4 所示。

表 4 海河网络边重要性的权值指标分布情况

Table 4 Distribution of weight index of edge importance in Haihe River Network			
等级	权值指标	边	边数/条
1	5.000×10^{-3}	(b13, S7)	1
2	4.100×10^{-3}	(H7, S2)	1
3	$(3.300\sim3.500)\times10^{-3}$	(R6, H7), (H2, S9)	2
4	2.100×10^{-3}	(b9, H2)	1
5	$(1.100\sim1.800)\times10^{-3}$	(S1, b13), (H6, H5), (P10, P1)	3
6	$(1.000\sim9.000)\times10^{-4}$	(R7, b9), (b12, S1), (b11, b12), (S9, S1), (S7, o1)等	43
7	$(1.096\sim9.683)\times10^{-5}$	(P5, P6), (s8, R10), (b3, R8), (s23, o1), (s14, b10)等	22
8	$(4.168^6\sim8.049)\times10^{-6}$	(s17, b7), (s12, b11), (s24, R3), (s15, b6), (s19, b8)	5

3.2.3 2 种指标对比分析

边的介数指标主要考虑边的信息传递能力,具有普遍通用性。由于只需要考虑网络的最短路径,介数指标的计算量相对较少。在海河网络中,介数指标位于 $[3.000\times10^{-4},4.510\times10^{-2}]$ 区间,且多数边的介数指标不相等。

边的权值指标具体考虑了河道的长度、流量、最短路径等因素,更能表征河流网络的边重要性。在海河网络中,权值指标位于 $[4.168\times10^{-6},5.000\times10^{-3}]$ 区间,且边的权值指标均不相等,因此边的权值指标能精确区分每条边的重要性。在权值指标中,海河干流具有最高的重要性,是整个海河网络的核心。而在介数指标中,海河干流位于等级 4,未能体现其重要性,与实际情况不相符。因此,权值指标更能准确找到河流网络的重要边。但边的权值指标的计算量较大。

4 结 语

引入复杂网络理论,建立了不含工程节点的海河网络模型 I 和含有工程节点的海河网络模型 II。

针对河流网络节点的特殊性,提出了河流网络节点重要性的权值指标。研究了海河网络的节点重要性,证明权值指标综合考虑了节点的自身性质、信息传递能力和节点间的距离等因素,能更准确地找到河流网络的重要节点。针对河流网络边的特殊性,提出了河流网络边重要性的权值指标。通过计算海河网络的边重要性,证明边的权值指标考虑更全面,更能表征河流网络的边重要性。

本文只是在所建河网模型的基础上研究河流网络的节点及边的重要性指标,若能将其应用到河流网络模型的简化研究中,删除河流网络中相对不重要的节点和边,从而突出相对重要的节点和边,有利于保护、监控河流网络中的重要信息,则更能体现其应用价值。

参考文献:

- [1] 倪向萍, 梅生伟, 张雪敏. 基于复杂网络理论的输电线路脆弱度评估方法[J]. 电力系统自动化, 2008, 32(4): 1-5. (NI Xiangping, MEI Shengwei, ZHANG Xuemin. Transmission lines' vulnerability assessment based on complex network theory [J]. Automation of Electric Power Systems, 2008, 32(4): 1-5. (in Chinese))
- [2] 赵月, 杜文, 陈爽. 复杂网络理论在城市交通网络分析中的应用[J]. 城市交通, 2009, 7(1): 57-65. (ZHAO Yue, DU Wen, CHEN Shuang. Application of complex network theory to urban transportation network analysis [J]. Urban Transport of China, 2009, 7(1): 57-65. (in Chinese))
- [3] SPORNS O. The human connectome: a complex network[J]. Ann N Y Acad Sci, 2011, 1224: 109-125.
- [4] 田占伟, 隋瑒. 基于复杂网络理论的微博信息传播实证分析[J]. 图书情报工作, 2012, 56(8): 42-46. (TIAN Zhanwei, SUI Yang. The empirical analysis of micro-blog information flow based on complex network theory[J]. Library and Information Service, 2012, 56(8): 42-46. (in Chinese))
- [5] RINALDO A, RODRIGUEZ-ITURBE I, RIGON R, et al. Self-organized fractal river networks[J]. Physical Review Letters, 1993, 70(6): 822-825.
- [6] CIEPLAK M, GIACOMETTI A, MARITAN A, et al. Models of fractal River basins[J]. Journal of Statistical Physics, 1998, 91(1): 1-15.
- [7] 何大韧, 刘宗华, 汪秉宏. 复杂系统与复杂网络[M]. 北京: 高等教育出版社, 2009.
- [8] RODRIGUEZ-ITURBE I, RINALDO A, RIGON R, et al. Fractal structures as least energy dissipation patterns: the case of river networks[J]. Geophysical Research Letters, 1992(5): 2854-2860.
- [9] RINALDO A, RODRIGUEZ-ITURBE I, RIGON R, et al. Self-organized fractal river networks[J]. Physical Review Letters, 1993, 70(3): 822-825.
- [10] 沈丹, 许田, 苏蓓蓓, 等. 长江流域的河流网络及其统计描述[C]//方锦清. 第二届全国复杂动态网络学术论坛论文集. 北京: 中国高等科学技术中心, 2005: 662-664.
- [11] 许田. 对一些复杂网络的统计描述与性质的研究[D]. 扬州: 扬州大学, 2004.
- [12] 汪小帆, 李翔, 陈关荣. 复杂网络理论及其应用[M]. 北京: 清华大学出版社, 2006.
- [13] HE Jinliao, ZONG Yueguang, YANG Wei. The analysis of water network of Beijing-Tianjin region based on the complex network theory[C]//LI Xia, LIU Lin. 16th International Conference on Geoinformatics and the Joint Conference. Guangzhou: Sun Yat-sen University Press, 2008.
- [14] 全国水力资源复查工作领导小组. 中华人民共和国水力资源复查成果(2003年)第一卷[M]. 北京: 中国电力出版社, 2004.
- [15] GIRVAN M, NEWMAN M E J. Community structure in social and biological networks[J]. Proc Natl Acad Sci, 2002, 99: 7821-7826.