利用重要性贡献矩阵确定通信网中最重要节点

赵毅寰 王祖林 郑 晶 郭旭静

(北京航空航天大学 电子信息工程学院, 北京 100191)

摘 要: 针对已有的节点重要性评价方法,基于节点的移除可能导致网络拓扑结构变化的不足,提出了一种利用节点间关联特性的通信网络节点重要性评价方法.该方法定义的节点重要性贡献矩阵(NICM,Node Importance Contribution Matrix)考虑了网络中不同节点间的联接关系对节点重要性的影响,每个节点对其相邻节点重要程度的贡献与该节点的度有关,节点的初始重要性设为该节点的介数.利用该算法对典型网络的节点重要性进行了分析,并且与移除节点的方法进行了对比.实验结果表明,所提出的算法无须考虑网络拓扑结构的变化,结果精确并能正确评价节点对网络资源的控制能力.

关 键 词: 通信网络;最重要节点;度;介数;拓扑

中图分类号: TN 915.02

文献标识码: A **文 章 编 号**: 1001-5965(2009)09-1076-04

Finding most vital node by node importance contribution matrix in communication netwoks

Zhao Yihuan Wang Zulin Zheng Jing Guo Xujing

(School of Electronics and Information Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100191, China)

Abstract: Node importance evaluation based on removal of nodes and their incident links may change the topology of the network, even splits the network into several disconnected parts. The neighbor nodes' node importance contribution matrix (NICM) was defined and an improved node importance evaluation method was proposed. A node's importance value was initialized as its betweenness. Different node contributed importance value dissimilarly to a certain node, according to its degree and their position relationship. The algorithm was compared with the node-removal based method in the analysis of advanced research project agency(ARPA) net topology. The results show that this algorithm accurately represents the node importance, with less calculation workload and no change of the net topology. Nodes' ability to control the network recourse is also involved in this method.

Key words: communication network;most vital node;degree;betweenness;topology

可靠性是通信网络运行质量的一个重要指标,在网络的设计和维护过程中具有十分重要的意义.节点的重要性研究源于故障管理的需要.网络攻防的防御方对于网络中的重要节点,在其硬件质量和安全设置等方面都需要给予特别的关注;进攻方则需要以重要节点作为攻击目标以取得更理想的攻击效果.

确定网络中最短路径上的最重要节点(MVN,Most Vital Node)的问题最初由 Corley 和 Sha[1]在 1982 年提出,他们定义两个节点间的最重要节点是移除后使这两个节点间的最短距离增加最多的那个节点,这种方法只适用于两个端点的情况,无法评价全网范围的的节点重要性.文献[2]利用跳面的概念简化了由于网络中存在大

收稿日期: 2008-07-14

基金项目: "十一五"国防预研基金资助项目(41101040603)

作者简介: 赵毅寰(1982—),男,河南南阳人,博士生,zhaoyihuan@gmail.com.

量可用路由而导致的高算法复杂度,但忽略了跳面内的路由细节,影响了评价的精度.文献[3]通过比较移除节点后生成数的数目变化来确定最重要节点,对文献[2]的算法进行了改进.文献[4]利用节点收缩的方法,删除待分析节点相邻的节点,简化了计算复杂度,但上述方法都没有考虑移除节点后导致网络拓扑变化时的节点重要性.文献[5]讨论了由于移除节点而导致网络被分割的情况,提出了基于网络性能梯度的节点重要性评价方法,具有更广泛的适用范围,但是网络被切割的临界点难以判断,梯度权重的选择缺少一般性的方法.

上述方法均利用移除节点,比较网络性能变化的方法来确定最重要节点,这种各个节点单独分析的方法需要解决由于节点移除导致的网络拓扑结构变化甚至被分割的问题.本文将各个节点联系起来考虑,从分析节点对网络资源控制能力的角度,提出了一种基于节点关联性的重要节点评价方法,解决了这个问题.

1　建模及相关概念

设图 $G=(V,E)$ 是一个无自环的无向网络,其中 $V=\{v_1,v_2,\cdots,v_N\}$ 是网络中所有节点的集合;$E=\{e_1,e_2,\cdots,e_n\}$ 且 $E\subseteq V\times V$ 是节点间通路的集合.

定义 1　距离(distance).两节点 v_i 和 v_j 之间的距离 d_{ij} 是这两个节点之间的最短路径.如果 v_i 和 v_j 之间不存在路径,则 $d_{ij}\to\infty$.

定义 2　网络直径(net diameter).网络中所有节点对距离的平均值.网络直径反映了通信网络的性能,直径越大表示单个业务的平均连接长度越长,单个业务占用的网络资源越多,在网络规模一定时,网络能够提供的服务就越少;另一方面,网络规模一定时,网络直径越大,表示可替换的路由越少,网络的抗毁性就越差.

定义 3　介数(betweenness).如果一对节点间的最短路径(距离)有 g_{jk} 条,其中有 $g_{jk}(i)$ 条经过节点 v_i,那么节点 v_i 对这一对节点的介数的贡献为 $g_{jk}(i)/g_{jk}$.把节点 v_i 对所有节点对的贡献累加起来再除以节点对的总数,就可以得到节点 v_i 的介数 B_i.即

$$B_i=\frac{\sum_{j}^{n}\sum_{k}^{n}\frac{g_{jk}(i)}{g_{jk}}}{n(n-1)/2}\quad j\neq k\neq i\quad j<k\quad(1)$$

其中,n 是网络节点的总数.一个节点的介数反映了它在完成网络通信任务中提供最短可用路由的能力,可以用来衡量此节点对网络资源的控制能力.

2　基于关联性的重要性评价方法

网络是实现终端间互相通信的统一整体,表现在节点之间是相互关联的.某个节点的添加或者移除,都可能导致其他节点间可用路由的增减和节点负荷的变化,这也正是目前多数节点重要性评价方法的基础.因此,任何互相连通的节点间都存在一定的节点重要性依赖关系,而最直接也是最重要的依赖关系存在于相邻(直接相连)节点之间.

2.1　用节点的度构建节点重要性关联矩阵

节点的度指的是与之相邻的节点的数目,目前普遍认为[1-5]节点的重要程度与节点的度有很大的关系,所以,节点之间重要性的依赖关系也可以很方便地用度来确定.

孤立的节点彼此不存在重要性的依赖关系.一旦节点间建立起连接,某个节点的存在可能会导致另一个节点负荷或可用路由的变化,也就影响到其余节点的重要性.这种节点重要性的依赖关系可以用下面的例子说明.

如图 1 所示,环状网由于其对称性,所有节点间的依赖关系相同.在其中心添加一个节点 v_i,并与所有节点相连接,可以直观地看出节点间的可用路由从唯一一条增加到多条,原有节点重要性"转移"到了中心节点上,称这种节点重要性的变化为节点 v_j 对节点 v_i 重要性的贡献.也就是说,除了节点所处位置的贡献外,节点的重要性还可以认为来自其相邻节点的贡献.通过节点之间的传递,就形成一个重要性依赖关系的拓扑,它的结构来自实际的网络拓扑.这种节点间的重要性依赖拓扑,可以用一种代表重要性关联关系的矩阵来表示.

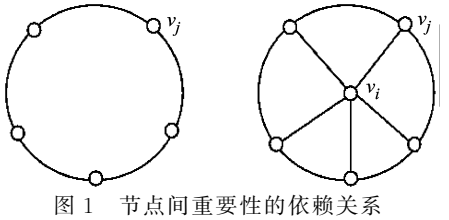

图 1　节点间重要性的依赖关系

定义 4　节点重要性贡献矩阵(NICM,Node Impor-tance Contribution Matrix).在 n 节点无自环无向网络中,若节点 v_i 的度为 D_j,则 v_i 将自身重要性的 $1/D_j$ 贡献给它的每一个相邻节点.将所有节点对其相邻节点的重要性贡献用矩阵的

形式表示出来，就形成了节点重要性贡献矩阵 $\boldsymbol{H}_{\mathrm{NICM}}$，记为

$$\boldsymbol{H}_{\mathrm{NICM}}=\begin{array}{c} \\ v_1 \\ v_2 \\ \vdots \\ v_n \end{array}\begin{array}{c} \begin{array}{cccc} v_1 & v_2 & \cdots & v_n \end{array} \\ \begin{bmatrix} I_1 & \delta_{12}/D_2 & \cdots & \delta_{1n}/D_n \\ \delta_{21}/D_1 & I_2 & \cdots & \delta_{2n}/D_n \\ \vdots & \vdots & \cdots & \vdots \\ \delta_{n1}/D_1 & \delta_{n2}/D_2 & \cdots & I_n \end{bmatrix} \end{array} \tag{2}$$

其中，贡献度分配参数 δ_{ij} 在节点 v_i 和 v_j 是相邻节点的时候取值为1，否则取值为0；对角线上的数 I_k 表示节点 v_k 的重要性的初始值. 下面讨论节点重要性初值的选择.

2.2 NICM中节点重要性初值的选取

如定义3所述，一个节点的介数反映了此节点为网络提供最短可用路由的能力，而且很多情况下节点对之间的最短路由是唯一的. 所以节点介数越大，移除此节点后网络直径变大的可能性就越大，即网络性能下降. 可见，节点的介数一定程度上反映了节点的重要程度.

从另一个角度考虑，节点的重要性取决于两个因素：节点的位置和节点的度[4]. NICM利用节点的度来构建节点之间的重要性关联，而节点的介数则是由其在网络拓扑中所处的位置决定的，即节点的介数反映了节点的位置信息. 因此，选取节点的介数作为NICM中节点重要性的初值. 结合了介数信息的NICM可以表示为

$$\boldsymbol{H}_{\mathrm{NICM}}=\begin{array}{c} \\ v_1 \\ v_2 \\ \vdots \\ v_n \end{array}\begin{array}{c} \begin{array}{cccc} v_1 & v_2 & \cdots & v_n \end{array} \\ \begin{bmatrix} B_1 & \delta_{12}/D_2 & \cdots & \delta_{1n}/D_n \\ \delta_{21}/D_1 & B_2 & \cdots & \delta_{2n}/D_n \\ \vdots & \vdots & \cdots & \vdots \\ \delta_{n1}/D_1 & \delta_{n2}/D_2 & \cdots & B_n \end{bmatrix} \end{array} \tag{3}$$

可以看出，一个节点对其相邻节点的重要性贡献与此节点本身的重要性和度的大小有关.

2.3 节点重要性评价算法

现有的节点重要性评价方法主要是比较移除节点前后网络性能的变化，对由于节点的移除可能造成的网络被分割的情况需要区别对待，且难以表达. 将节点重要性贡献矩阵和节点的介数结合起来评价通信网节点的重要性，相当于同时利用了节点的位置信息和度的信息，可以获得比较精确的评价结果. 提出的NICM结合节点介数的节点重要性评价算法的步骤如下：

Begin

1) 确定节点重要性贡献矩阵

For $i=1$ to n

{确定节点 v_i 的度和其相邻节点；

按照定义填入NICM中；}

2) 确定节点重要性初值

For $i=1$ to n

{确定节点 v_i 的介数；

按照定义填入NICM中；}

3) 计算每个节点对其他节点的重要性贡献

For $i=1$ to n

{将第 i 列上除了第 i 行的数都乘以初始值 $\boldsymbol{H}_{\mathrm{NICM}}(i,i)$；}

4) 计算每个节点的重要性

For $i=1$ to n

{将第 i 行上的所有数相加，输出 X_i；}

End

上述第4)步的输出 X_i，$i=1,2,\cdots,n$ 即依次为网络 n 个节点的重要性评价结果.

3 实验结果分析

下面分别针对简单网络和复杂网络，将上述算法与已有算法的评价结果进行对比，来说明其有效性.

图2表示的网络共有5个节点，它结合了节点介数的NICM具有如下形式：

$$\boldsymbol{H}_1=\begin{bmatrix} 0 & 0.3333 & 0.25 & 0 & 0 \\ 0.5 & 0.05 & 0.25 & 0.3333 & 0 \\ 0.5 & 0.3333 & 0.2 & 0.3333 & 0.5 \\ 0 & 0.3333 & 0.25 & 0.05 & 0.5 \\ 0 & 0 & 0.25 & 0.3333 & 0 \end{bmatrix} \tag{4}$$

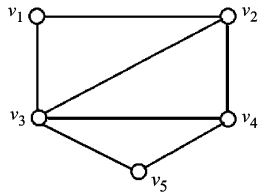

图2 一个简单的网络

表1是按照算法步骤计算出节点重要性评价结果 X 与文献[3]的评价结果 γ 对比的情况.

表1 图2所示网络的节点重要性

节点	本文		文献[3]	
	B	X	τ	γ
v_1	0	0.2167	8	0.6190
v_2	0.05	0.2667	3	0.8571
v_3	0.2	0.5334	1	0.9524
v_4	0.05	0.2667	3	0.8571
v_5	0	0.2167	8	0.6190

对比可知，对图2所示的简单网络，本文算法

的节点重要性评价结果同文献[3]的评价结果是一致的.本文算法不涉及矩阵乘法以及归一化等运算;简单网络拓扑的节点介数的确定也比较方便,因此运算量相对较小.

利用图 3 对复杂网络的节点重要性评价算法进行分析.这是美国的 ARPA(Advanced Resarch Project Agency)网络拓扑,由 21 个节点和 23 条链路组成.ARPA 拓扑是目前分析网络节点重要性时普遍使用的干线网络拓扑,其网络平均度值在 2～3 之间,大部分节点的度值为 2.

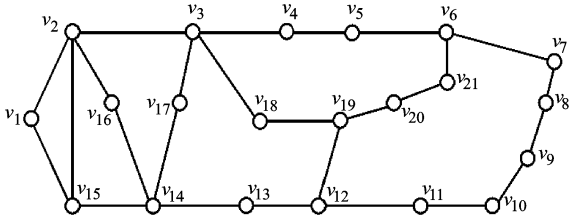

图 3　ARPA 网络拓扑

表 2 是本文提出的算法以及文献[5]和文献[3]所述方法确定的节点重要性评价结果.3 种算法得出的节点重要性排序都略有差异,是因为各自的判断侧重点不同.本文算法评价的依据是各个节点对网络资源的控制能力;而文献[5]和文献[3]分别依据的是移除节点后,网络直径的变化和生成树的数目变化.本文算法得出的最重要节点是 v_3,同文献[3]得出的结论是一致的.

表 2　ARPA 网络拓扑的节点重要性

节点	本文		文献[5]		文献[3]
	B	X	g_c	g_d	γ
v_1	0	0.035 5	0	−0.045 0	0.626 2
v_2	0.089 7	0.177 0	0.05	0.073 2	0.972 1
v_3	0.280 2	0.452 6	0.05	0.627 2	0.993 0
v_4	0.153 6	0.294 8	0	0.307 6	0.838 7
v_5	0.142 5	0.277 6	0	0.228 7	0.838 7
v_6	0.175 0	0.329 4	0	0.484 0	0.983 6
v_7	0.099 6	0.189 9	0	0.189 2	0.879 7
v_8	0.063 9	0.145 8	0	0.052 4	0.879 7
v_9	0.064 3	0.139 9	0	0.031 3	0.879 7
v_{10}	0.087 3	0.192 1	0	0.126 1	0.879 7
v_{11}	0.145 2	0.268 4	0	0.336 6	0.879 7
v_{12}	0.238 5	0.439 3	0	0.712 9	0.978 0
v_{13}	0.127 0	0.246 6	0	0.097 1	0.805 1
v_{14}	0.160 3	0.262 5	0	0.289 2	0.986 4
v_{15}	0.039 3	0.101 8	0	0.055 0	0.878 7
v_{16}	0.008 3	0.070 8	0	−0.039 7	0.663 9
v_{17}	0.042 9	0.153 0	0	0.023 4	0.697 7
v_{18}	0.103 6	0.238 3	0	0.207 6	0.770 1
v_{19}	0.194 0	0.363 0	0	0.476 1	0.967 1
v_{20}	0.075 4	0.173 4	0	0.178 7	0.827 9
v_{21}	0.066 7	0.162 7	0	0.099 7	0.827 9

注:g_c—网络连通系数梯度;g_d—网络直径梯度.

从表 2 还可以看出,文献[3]认为节点 v_7 到 v_{11} 的重要性是相同的,而根据本文的算法,这 5 个节点的重要程度是有区别的,这一点同文献[5]的结论是一致的.也就是说,本文的算法和基于网络性能变化梯度的算法都能够分辨出 v_7 到 v_{11} 这 5 个节点重要性的差别,比文献[3]的方法具有更高的评价精度;同时,利用节点关联性的算法不会导致网络结构变化,不存在网络分割临界的判断和描述量权重的选择问题[5].所以,本文提出的算法在提高节点重要性评价精度的情况下,具有更好的可操作性.

4　结 束 语

通信网中节点的重要性表征了节点发生故障的情况下对网络性能影响的程度.本文基于互联节点间的相互关系,定义了节点重要性贡献矩阵,在此基础上提出了 NICM 结合介数的节点重要性评价算法.它以介数作为初始值,侧重于节点对网络资源的控制能力;无须改变网络拓扑结构,运算简单且评价结果精确.NICM 定义了相邻节点间的重要性依赖关系,如何将它进行扩展,使之能直接反映所有节点间的重要性依赖关系,是下一步要研究的内容.

参考文献 (References)

[1] Corley H W, Sha D Y. Most vital links and nodes in weighted networks [J]. Oper Res Letters, 1982(1): 157−160

[2] 郭伟. 野战地域通信网可靠性的评价方法[J]. 电子学报, 2000, 28(1): 3−6
Guo Wei. Reliability evaluation method of tactical communication network [J]. Journal of Electronics, 2000, 28(1): 3−6 (in Chinese)

[3] 陈勇,胡爱群,胡啸. 通信网中节点重要性的评价方法[J]. 通信学报, 2004,25(8): 129−134
Chen Yong, Hu Aiqun, Hu Xiao, Evaluation method for node importance in communication networks [J]. Journal of China Institute of Communications, 2004,25(8): 129−134 (in Chinese)

[4] Wu Jun, Tan Yuejin. Finding the most vital node by node contraction in communication networks [C]//Communications, Circuits and Systerms, Proceedings of IEEE International Conference. 2005,2:1282−1286

[5] 余新,李艳和,郑小平,等. 基于网络性能变化梯度的通信网络节点重要程度评价方法[J]. 清华大学学报:自然科学版, 2008,48(4): 541−544
Yu Xin, Li Yanhe, Zheng Xiaoping, et al. Node importance evaluation based on communication network performance grads [J]. J Tsinghua Univ:Sci & Tech, 2008, 48(4): 541−544(in Chinese)

[6] Albert R, Jeong H, Barabasi A L. Error and attack tolerance of complex networks[J]. Nature, 2000, 406: 378−382

(责任编辑:娄　嘉)