物理学报 Acta Physica Sinica
Chinese Physical Society
Institute of Physics, CAS

推荐重要节点部署防御策略的优化模型

杨雄　黄德才　张子柯

Recommendation of important nodes in deployment optimization model of defense strategy

Yang Xiong　Huang De-Cai　Zhang Zi-Ke

引用信息 Citation: Acta Physica Sinica, 64, 050502 (2015)　DOI: 10.7498/aps.64.050502
在线阅读 View online: http://dx.doi.org/10.7498/aps.64.050502
当期内容 View table of contents: http://wulixb.iphy.ac.cn/CN/Y2015/V64/I5

推荐重要节点部署防御策略的优化模型*

杨雄[1)2)] 黄德才[1)†] 张子柯[3)]

1) (浙江工业大学计算机科学与技术学院, 杭州 310023)

2) (常州工学院计算机信息工程学院通信工程系, 常州 213002)

3) (杭州师范大学阿里巴巴复杂科学研究中心, 杭州 311121)

(2014 年 7 月 26 日收到; 2014 年 10 月 10 日收到修改稿)

当前网络安全防御策略集中部署于高连接度节点主要有 2 个方面的不足: 一是高连接度节点在很多场合中并不是网络通信的骨干节点; 二是该类节点对信息的转发和传播并非总是最有效的. 针对以上传统部署策略的不足, 改进了恶意病毒程序传播的离散扩散模型并采用中间路径跳数来衡量网络节点的重要程度, 提出了基于介数中心控制力和接近中心控制力模型的重要节点优先推荐部署技术. 实验结果显示具有高介数中心控制力和低接近中心控制力的节点相对于传统的高连接度节点无论在无标度网络还是小世界网络均能够对恶意病毒程序的疫情扩散和早期传播速度起到更加有效的抑制作用, 同时验证了网络分簇聚类行为产生的簇团特性也将对恶意程序的传播起到一定的负面影响.

关键词: 高连接度, 介数中心, 接近中心, 分簇

PACS: 05.10.–a, 02.30.Oz **DOI:** 10.7498/aps.64.050502

1 引 言

2014 年文献 [1] 指出以 email 蠕虫和社会网络蠕虫病毒为代表的新一代网络恶意程序将对互联网产生严重的威胁. 这类新型恶意程序通常都对网络拓扑信息非常敏感, 比如他们通常都会依赖于被攻陷主机内部的 email 地址簿信息或者社区通讯工具内含有的好友信息来定位攻击对象. 这种智能化的攻击策略相对于随机扫描大量地址空间进行传播扩散的传统感染方式显得更加高效, 此类新型恶意程序通常能够快速定位攻击对象并快速感染攻击目标从而使得感染规模迅速扩大. 更重要的是, 借助于社会工程学技术, 该类病毒通过诸如好友发送的电子邮件或者文件传输进行攻击, 而被攻击目标通常出于信任好友而无法识别恶意代码导致感染.

为了消除这种对网络拓扑信息非常敏感的智能网络病毒, 同时有效控制和抑制病毒的爆发, Zou 等 [2] 从离散的角度对 email 蠕虫的传播动力学进行了建模并给出了防御策略; 2010 年 Fan 等 [3] 首次提出了 P2P 蠕虫的扩散模型并利用少数 peer 端点进行抑制来实现大多数终端的防御策略. 2011 年, Yan 等 [4] 针对在线社会网络中的恶意软件进行了全面的动力学和防御策略研究, 对初始感染节点数量、用户点击概率、社会结构和个体行为模式等参数在社会网络恶意软件传播的影响力方面进行了仿真研究; Xing 等 [5] 分析了多个小世界网络组成的复杂网络模型中谣言的传播动力学模型, 该模型借鉴了传染病传播机理并能反映真实世界中社团之间的有组织通信方式. 2012 年 Song 等 [6] 和 Lu 等 [7] 利用平局场理论提出了无标度网络中的恶意软件传播模型, 该类模型对传统传染病的脆弱-感染-免疫模型 (susceptible-infected-recovered, SIR) 进行了改进, 考虑了节点在收到邻居节点被感染的预警信息后采取免疫措施的情形. 同年, 文献 [8]

* 浙江省自然科学基金 (批准号: LY12A05003, LQ13F030015) 和江苏省高校自然科学基金 (批准号: 13KJD520001) 资助的课题.

† 通信作者. E-mail: hdc@zjut.edu.cn

从另一个角度阐述了网络中重要传播节点的定位方法, 该文认为处于网络中核心的中央节点对信息的传播具有更强的影响力, 同时也研究了当多个传播源同时进行信息传递时, 传播源节点之间的距离也将会对最终的传播效果起到举足轻重的作用. 2011 年Schneider 等[9] 提出了一种只需要低成本的网络结构改变就可以高效提升不同网络强壮度的方法, 实验证明这种方法不仅可以提高网络的强壮度还可以用来设计高效和稳健的网络系统. 2014 年Zhao 等[10] 研究了关于多层网络的新型多途径病毒传播模型, 分析了 Commwarrior 这种能够同时根据手机电话簿构建短信网络和根据地理位置构建蓝牙网络进行传播的恶意程序传播机理; 提出了该型网络中的关键节点是在不同网络层次具有度-度 (连接度) 相关性较大的节点; 指出这类节点会促进网络变成异构网络从而导致病毒传播门限值越小, 越容易早期爆发 (因为高连接度节点的存在), 但是在稳定阶段相对于同构网络则会产生较小的感染规模 (因为低连接度节点很难被感染). 文献 [11] 研究了恶意程序在节点连接度差异性网络环境中的传播模型; 同时文献 [12, 13] 创新性地指出无标度网络中具有高连接度的骨干节点和小世界网络中具有强簇团性的中心节点往往能够在蠕虫病毒爆发的早期阶段加速疫情的传播. 但是如何高效选择网络的防御结构和防御策略的部署位置仍然是个尚未解决的研究难题, 研究领域普遍的观点就是防御策略通常优先部署在具有高连接度的骨干节点, 然而这些研究成果通过在高连接度的骨干节点进行防御部署是否同样能够对拓扑信息敏感的智能恶意程序起到抑制作用仍然是个未解的课题.

为了直观地理解节点在不同网络拓扑中的地位, 图 1 给出了两种特殊的网络结构. 从图 1 (a) 中可以看到网络结构被分成左右两个组, 他们通过类似于桥接的关键点进行彼此互通. 在该图中节点 v_3 和 v_7 具有较大的连接度, 他们的连接度均为 5; 而 v_6 和 v_{12} 尽管连接度为 2 但是他们却对整个网络的互通起到了决定性作用, 具有较高的介数中心控制力, 因此在进行防御隔离部署策略时节点 v_6 和 v_{12} 应该具有更高的优先部署权.

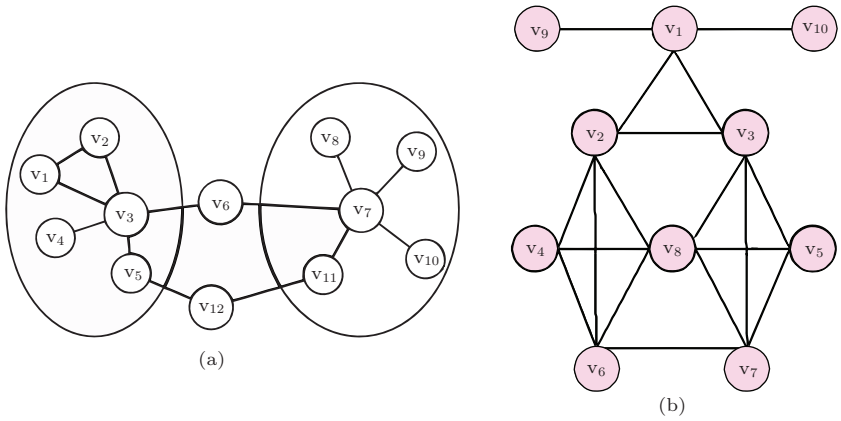


图 1 网络节点角色示意图 (a) 介数中心节点示意图; (b) 接近中心节点示意图

图 1 (b) 显示了接近中心节点的概念, 在该图中节点 v_8 具有最大的连接度, 连接度为 6; v_2 和 v_3 连接度相对于 v_8 较小, 连接度均为 5; 但是从 v_2 和 v_3 出发却具有到其他节点最短的路径, 也就是说从这两点出发能够比从 v_8 出发去往其他节点速度更快, 因此从攻击者角度而言选择 v_2 和 v_3 作为蠕虫病毒的感染源更具吸引力.

从上面的讨论中可以看到具有高连接度的节点并非总是网络的骨干节点, 选择这些节点进行防御部署并不一定能够取得较好的效果. 值得说明的是, 本文提出的基于中间跳数的部署优化模型是针对单层网络的, 从介数中心和接近中心这两个节点的本质属性出发研究针对拓扑信息敏感的智能恶意程序的防御优化部署策略相信也能对将来多层网络传播的病毒防御起到一定的推广效应. 研究结果显示对具有高连接度的节点部署防御并非总是有效的防御策略, 仿真实验结果表明基于介数中心和接近中心的重要节点推荐部署防御策略能够有效优化防御的效果.

2 基于中间跳数的蠕虫病毒传播模型

利用跳数可以有效研究信息在网络中的中间传播过程和状态, 比如P2P 的种子搜索过程就是采用了跳数的思想, 为了尽可能得到有用的种子, 搜索过程大多采用广播的形式进行穷举搜索, 但是这样的方式通常会产生泛洪流量, 因此采用了限制搜索时间和跳数范围的技术; 基于随机漫步的搜索则为了克服泛洪风暴的不利因素在选择下一步搜索节点时, 随机选择邻居节点中的一个进行转发. 本文为了模型建模的方便假设了病毒传播过程中根据攻击目标周边所有的邻居状态来决定目标被感染的概率, 并没有限制信息在网络中跳数的范围和时间. 为了评估在推荐重要节点实施蠕虫病毒防御安全策略的抑制效果, 本文按照以下5个步骤对拓扑信息敏感的智能病毒扩散模型进行离散建模.

1) 首先一个被感染节点可以通过攻击策略向外传送恶意代码, 而潜在的攻击目标则可能因为漏洞或者社会工程学的原因感染恶意代码并形成一个新的感染源. 通常情况下该种智能病毒从一个节点x依靠邮件地址列表或好友花名册定位攻击目标, 通过若干个中间节点到达目标节点y. 假设k表示从原点x到目标点y经历过的中间路径跳数, 并且$I_x(k)$表示节点x在经历过中间k跳后的感染状态. 如果节点x在经历中间k跳后变成感染状态, 则$I_x(k)=1$, 否则$I_x(k)=0$.

2) 为了表示节点之间的感染行为, 定义一个$N\times N$的初始邻居矩阵$\boldsymbol{T}$其中矩阵元素P_{xy}表示从两个邻居节点x到y的传染概率, 当一个节点状态处于脆弱可感染状态时, 那么该节点就可被其邻居节点感染, 如下式所示:

$$\boldsymbol{T}=\begin{bmatrix} P_{11} & & & \\ & P_{xy} & & \\ & & \cdots & \\ & & & P_{NN} \end{bmatrix},$$

$$P_{xy}\in[0,1],$$

$$P_{xy}=0(x=y\bigcup x\notin N_y), \tag{1}$$

(1)式表明了节点对自己不会感染或者当节点x不属于节点y的邻居节点时感染概率为0.

3) 将上面定义的P_{xy}概念进行延伸, 表示成恶意代码传播时在$k-1$跳前节点y的邻居节点x已经被感染且节点y尚未被感染处于脆弱状态的前提下, 第k跳传播后节点y被感染的概率. 因此给出

$$\begin{aligned} P_{xy}=&P(I_y(k)\\ &=1\,\Big|I_x(k-1)=1\bigcap I_y\,(k-1)=0\\ &\quad \text{or}\ I_y(k-1)=1). \end{aligned} \tag{2}$$

正如(2)式所示, 任意节点x在恶意代码传播k跳路径后被感染的状态$I_y(k)=1$只有在一种情况下发生: 当$k-1$跳前该节点已经被感染或者$k-1$跳前尚未感染但被邻居在第k跳后感染.

4) 因此基于(2)式可以得到任意节点x在恶意信息传播了k跳后的感染概率为

$$\begin{aligned} &P(I_x(k)=1)\\ =&P(I_x(k-1)=1)+P(I_x(k)=1,I_x(k-1)=0)\\ =&\,P(I_x(k-1)=1)+P(I_x(k-1)=0)\\ &\times P(I_x(k)=1\,|I_x(k-1)=0\,)\\ =&\,P(I_x(k-1)=1)+P(I_x(k-1)=0)\\ &\times P(I_x(k)=1\,|I_x(k-1)=0\,,I_y(k-1)=1,\\ &\quad \exists\ y\in N_x)\\ =&\,P(I_x(k-1)=1)+P(I_x(k-1)=0)\\ &\times\left\{1-\prod_{y=1}^{N}\left[1-P(I_x(k)=1\,|I_x(k-1)=0\,,\right.\right.\\ &\quad \left.\left.I_y(k-1)=1)\right]\right\}\\ =&\,P(I_x(k-1)=1)+P(I_x(k-1)=0)\\ &\times\left[1-\prod_{y=1}^{N}(1-P_{yx})\right]. \end{aligned} \tag{3}$$

需要说明的是在(3)式中当$k=1$时, 得到初始状态

$$\begin{aligned} &P(I_x(1)=1)\\ =&\,P(I_x(0)=1)+P(I_x(0)=0)\\ &\times\left[1-\prod_{y=1}^{N}(1-P_{yx})\right]\\ =&\,P(I_x(0)=1)+[1-P(I_x(0)=1)]\\ &\times\left[1-\prod_{y=1}^{N}(1-P_{yx})\right]\\ =&\,1-\prod_{y=1}^{N}(1-P_{yx}). \end{aligned} \tag{4}$$

在(4)式中借助于(1)式$P_{xx}=0$, 也就是说任何

节点都不会感染自己, 因此初始状态下 $P(I_x(0)=1)=0$, 而 P_{yx} 就是 (1) 式中的 $\boldsymbol{T}$ 矩阵初始元素值.

5) 通过以上的讨论可以利用恶意代码在网络中的传播中间跳数来计算出蠕虫病毒的感染模型, 令 $n(k)$ 表示经过中间 k 跳后的感染节点总数, 则离散感染模型为

$$n(k)=\sum_{x=1}^{N}P(I_x(k)=1). \tag{5}$$

3 防御部署优化模型

正如图 1 所示, 在高连接度节点部署防御策略并不一定能够取得理想的安全效果, 如何定位高效的防御策略部署位置仍然是一个重要的研究难题. 通常意义上一个理想的防御部署节点应该满足以下两个必要条件: 1) 一旦在重要节点部署防御策略后网络被感染的主机数量应该小于同样的策略部署在其他节点后网络被感染的主机数量; 2) 正如文献 [14] 所述, 在恶意代码扩散传播的早期阶段, 传播的速度很大程度上依赖于网络的规模, 一旦在重要节点部署防御策略后疫情的扩散速度应缓慢于同样安全策略部署在其他节点后疫情的传播速度. 因此为了定位部署防御策略的最佳节点, 本文将引入 Newman [15] 提出的介数中心节点和接近中心节点. 介数中心和接近中心的概念在很多研究中均被用来研究复杂网络节点的重要性, 通常情况下一个节点 x 的介数中心指任意两个节点连通彼此的中间路径中, 经过 x 节点的路径数比率; 节点 x 的接近中心指该节点与周边其他节点的路径之和. 介数中心和接近中心被认为是衡量网络节点重要性和中心性的重要指标, 本文从中间跳数的角度出发对这两个特性在网络安全防御定位的应用进行了研究. 为了模型建立的方便, 本文分别从接近中心和介数中心两个角度对模型进行定义.

3.1 基于接近中心控制力模型的优化部署

定义 $C_x(k)$ 为节点 x 在 k 跳路径内去往其他所有节点的平均距离. 将 $C_x(k)$ 表示为节点 x 的接近中心控制力, 该指标代表了从节点 x 出发去往其他节点的距离远近程度, 因此可以用该指标来衡量蠕虫病毒在一个给定节点向外扩散传播时的速度. 如果 $C_x(k)$ 值越大, 则表明节点 x 与外部其他节点的距离越远, 蠕虫病毒在该点向外扩散传播时速度越慢; 反之该参数值越小, 则说明节点 x 与其他节点距离越近, 病毒在该点向外传播时速度越快.

首先定义一个 $N\times N$ 的 k 阶矩阵 $\boldsymbol{T}^k$, 其中矩阵元素 P_{xy}^k 表示从源节点 x 到目标节点 y 经过中间 k 跳后的感染概率, 与 (1) 式不同的是这里的源节点 x 和目标节点 y 并非是物理相邻的邻居节点, 而是通过中间 k 个节点才能相互连通的任意两个节点. 定义如下:

$$\boldsymbol{T}^k=\begin{bmatrix} P_{11}^k & & & \\ & P_{xy}^k & & \\ & & \cdots & \\ & & & P_{NN}^k \end{bmatrix},$$

$$P_{xy}^k=1-\prod_{h=1}^{N}(1-P_{xh}^{k-1}P_{hy}), \tag{6}$$

式中的 P_{hy} 概念与 (1) 式相同, 表示互为邻居的两个节点 h 和 y 之间的感染概率, 而 P_{xh}^{k-1} 则表示从节点 x 到节点 h 经过 $k-1$ 跳中间路径的感染概率, 该参数可利用 (6) 式迭代计算得到.

从 (6) 式可以定义出任意节点 x 的接近中心控制力指标 $C_x(k)$ 和任意两个节点 x 和 y 在中间跳数 k 内的平均距离 d_{xy}^k, $C_x(k)$ 和 d_{xy}^k 的定义公式为

$$d_{xy}^k=\sum_{h=1}^{k}h\times P_{xy}^h, \tag{7}$$

$$C_x(k)=\frac{1}{N}\times\sum_{k=1}^{k}\sum_{y=1}^{N}d_{xy}^k. \tag{8}$$

从 (7), (8) 式可得对于任意节点 x 的接近中心控制力而言, $C_x(k)$ 的值越小则表明从节点 x 出发去往其他节点的距离越近, 攻击者选择该类节点作为攻击通常可以在有限的步骤内 (如 k 步) 将疫情的扩散进行最大化, 因此选择接近中心控制力值较小的节点进行防御布控可起到优化防御效果的作用.

3.2 基于介数中心控制力模型的优化部署

定义 $B_x(k)$ 为节点 x 在 k 跳路径内的介数中心控制力, 该指标表示了网络中任意两个节点进行通信需要经过节点 x 的 k 跳中间路径比率, 如定义 G_{ij} 为节点 i 到节点 j 的路径总数, G_{ij}^x 为节点 i 到节点 j 的路径中经过节点 x 的路径数, 则 $B_x(k)=G_{ij}^x/G_{ij}$, 因此可以利用该指标来衡量网络中节点的重要性程度. 如果 $B_x(k)$ 值越大, 则说明网络中任意两个

节点需经过 x 才可进行相互通信的路径占比越大, 节点 x 对整个网络通信的重要性越高; 反之该值越小, 则说明节点 x 对整个网络通信的重要性越低.

基于 (6) 式假设参数 μ_{xy}^{k} 表示在中间路径的 k 跳内从节点 x 到节点 y 的感染概率, 当 $k=0$ 时 $\mu_{xy}^{0}=P_{xy}$. 因此可以定义出任意节点 x 的介数中心控制力指标 $B_x(k)$ 和从节点 x 到节点 y 在中间跳数 k 内的平均感染概率如下所示:

$$\mu_{xy}^{k}=1-\prod_{h=1}^{k}(1-P_{xy}^{h}), \tag{9}$$

$$B_x(k)=\frac{1}{kN^2}\sum_{h=1}^{k}\left[\sum_{\text{start}=1}^{N}\sum_{\text{end}=1}^{N}(\mu_{\text{start},x}^{h}\times\mu_{x,\text{end}}^{k-h})\right]. \tag{10}$$

从 (9), (10) 式可以知道对于任意节点 x 的介数中心控制力而言, $B_x(k)$ 的值越大则表明节点 x 在网络中的地位越重要, 因此选择介数中心控制力值较大的节点进行防御布控可起到全局掌控疫情发展的作用.

4 实验分析

本实验环节采用 C 语言实现基于拓扑信息的智能病毒传播扩散行为, 在设计程序过程中每个节点包含一个好友节点标示符的邻接链表, 该链表表示了一旦节点被感染则恶意程序将会从中选择邻居好友进行目标攻击. 实验过程中为了不失一般性, 采用了与复杂社会网络最为接近的无标度网络拓扑, 网络规模为1000个节点, 为了统计数据的精准性, 我们对每个实验进行了100次计算以求得最终结果的平均值.

值得注意的是, 2013年Chen等[16]进一步挖掘了在网络局部结构中隐含的信息, 从而再次提高了有影响力节点挖掘的准确度, 该研究认为簇系数对于节点获取新的邻居具有消极负面影响, 因为簇系数越大说明社交的圈子越窄, 获取新节点的能力也就越差. 所以本实验也分别从两个角度阐述了在网络簇类系数大小两种情况下不同重要节点在恶意程序传播和抑制效果中扮演的不同角色. 为了显示网络结构和拓扑中具有高介数中心控制力及高接近中心控制力节点的分布, 图2和图3分别给出了相应的统计结果.

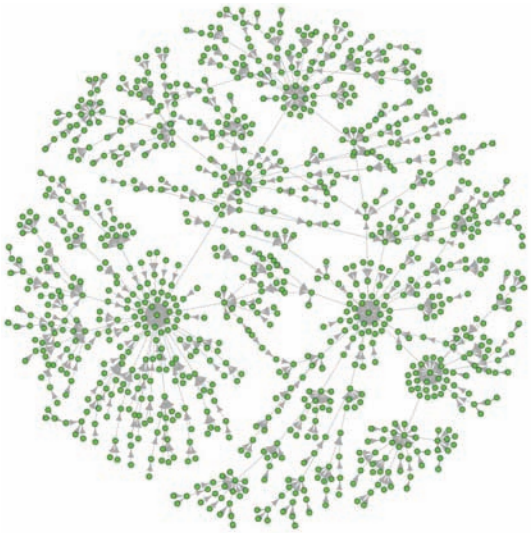


图2 (网刊彩色) 1000个节点的无标度网络

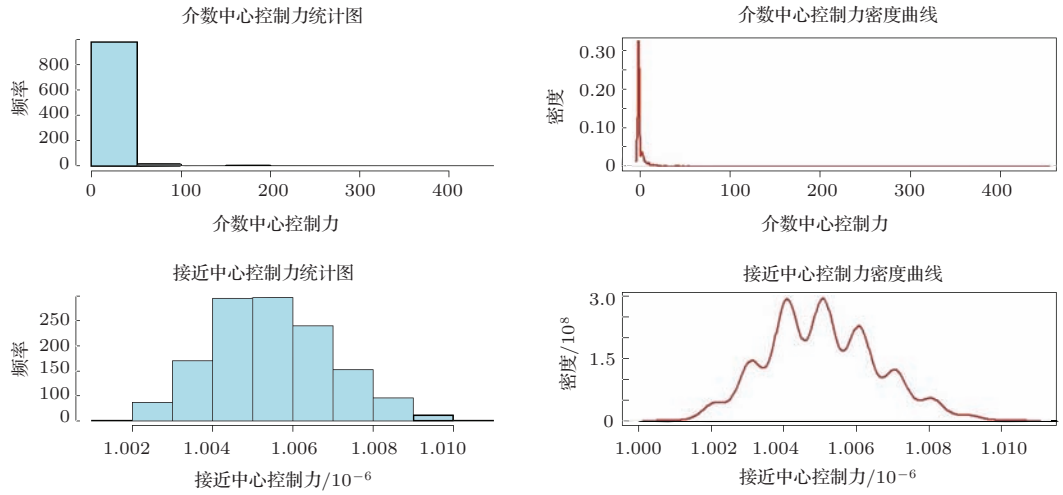


图3 (网刊彩色) 1000个节点的介数中心控制力和接近中心控制力统计图

需要说明的是图2表示了1000个节点的无标度网络, 但是该网络中节点都属于一个簇网, 并没有进行人为的簇网分割, 接下来的4.1节将重点分析图2所示网络中不同角色的节点对蠕虫病毒扩散和抑制的效果; 4.2节将重点分析同样规模的1000个节点无标度网络在进行不同尺寸的簇类分割后不同节点对蠕虫病毒扩散和抑制的效果. 图3显示了在无标度网络中节点介数中心控制力和接近中心控制力的统计分布, 在实验过程中取 $\text{Close} = \dfrac{1}{C_x}$ 是为了将介数中心控制力和接近中心控制力呈现一致方向性, 若两者的值越大表明节点在网络中的地位越重要, 即认为该节点是网络中重要的介数中心点同时也是越接近网络中心的节点. 图3可以看到绝大多数节点介数中心控制力都在50以下, 只有少数节点的介数中心控制力在150—200之间, 也就是说对于网络中能够对整个通信起到核心控制作用的节点呈现重尾分布的现象; 而接近中心控制力则呈现出类似于正态分布的现象, 有将近700个节点的接近中心控制力位于中等水平, 而只有不到50个节点的接近中心控制力是处于最高和最低水平, 换句话说大多数节点都位于距离网络中心中等远近的位置, 只有极少数节点处于网络的中心位置和末梢边缘位置.

4.1 重要节点对蠕虫病毒抑制和扩散的性能分析

通常情况下从攻击和防御两个角度出发一个理想的重要节点应该满足以下两个条件: 1) 从防御角度出发, 在该节点部署防御策略后被感染的主机数量应该小于相同防御策略部署在其他节点后网络被感染的主机数量; 2) 从攻击角度出发, 针对该节点的攻击将会导致疫情的传播和扩散加速. 实验中拓扑为图2所示的无标度网络, 分簇系数 $= 1$, 初始感染节点数 $= 10$, 我们分别选择具有高连接度、高介数中心控制力、低接近中心控制力的节点作为网络重要节点来分析他们对蠕虫病毒防御和攻击的性能效果. 实验结果如图4所示, 图4(a)中

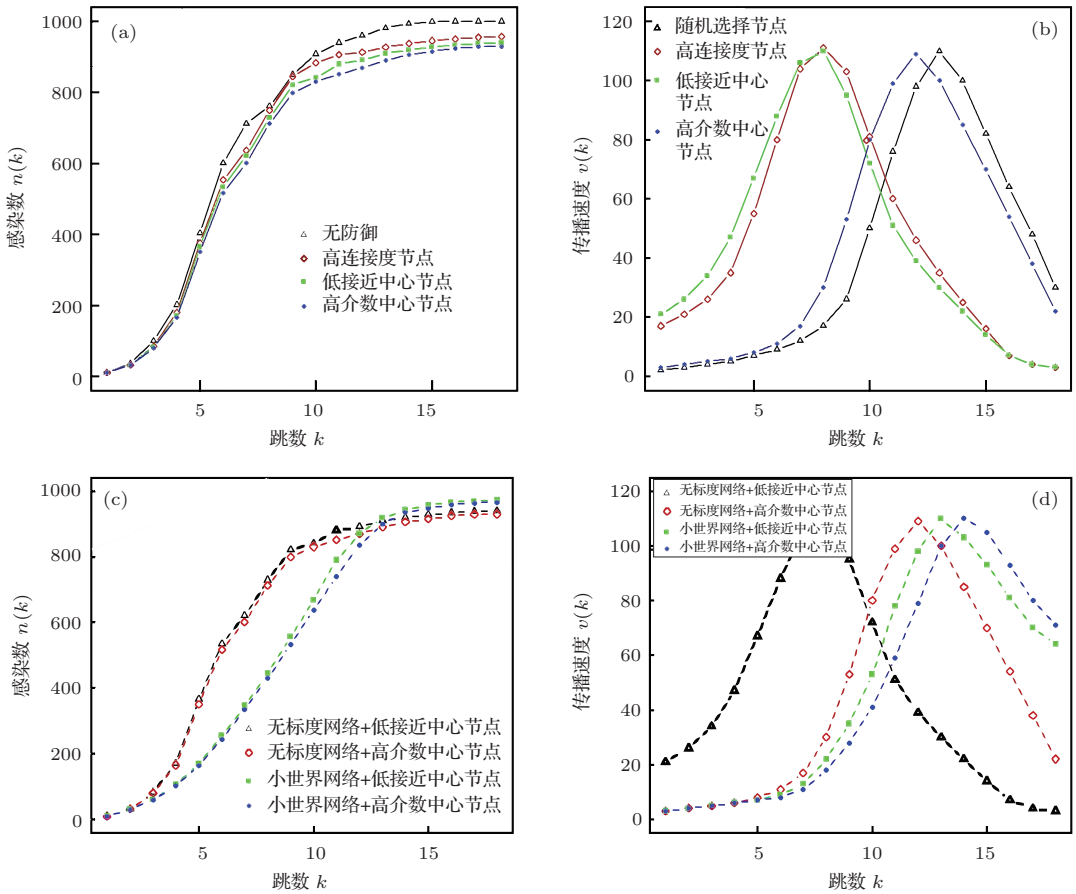
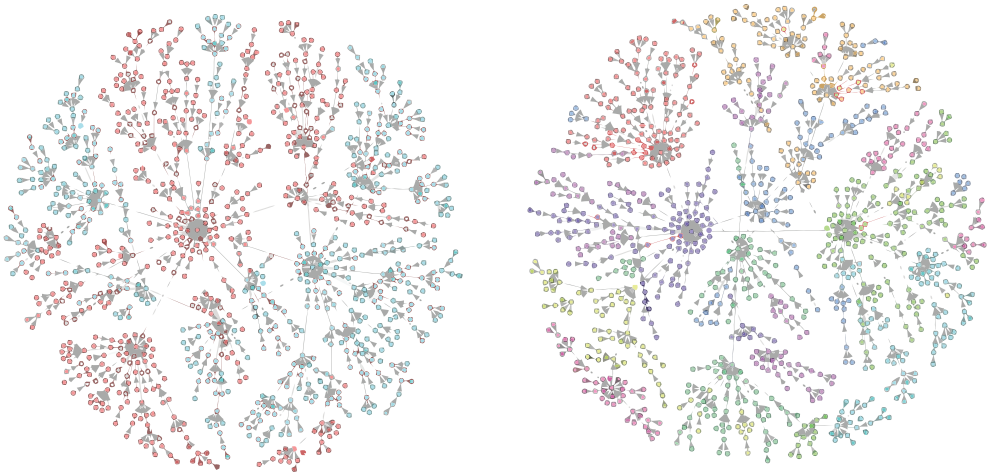


图4 (网刊彩色) 不同节点对恶意程序抑制和传播的效果分析 (a) 无标度网络中不同节点部署防御策略后的抑制效果; (b) 无标度网络中不同节点对病毒传播速度的影响效果; (c) 无标度网络和小世界网络中不同节点部署防御策略的抑制效果; (d) 无标度网络和小世界网络中不同节点对病毒传播速度的影响

显示了无标度网络中从防御角度出发在不同节点部署防御策略后的抑制效果, 非常明显的是在高介数中心控制力节点部署防御后病毒的感染效果取得了明显的抑制. 这种现象可以解释为在通信网络中那些必经之路的节点往往对蠕虫病毒的扩散起到了决定性的作用, 在该类节点部署安全策略可全面抑制恶意程序的扩散. 从攻击者角度出发一个重要节点应该是能够快速传播和扩散疫情的节点, 图 4 (b) 显示了无标度网络中具有低接近中心控制力的节点和高连接度的节点能够在蠕虫病毒爆发早期阶段最快速地传播恶意程序, 该现象可以解释成由于低接近中心节点和高连接度节点往往位于网络的中心区域, 与周边邻居距离较近或者有机会和更多的邻居节点通信, 导致了早期的传播速度非常的迅速. 因此与以往研究普遍认为高连接度节点应该优先部署安全策略不同, 具有高介数中心和低接近中心的节点更应该被认为优先部署防御策略. 为了验证提出的方案具有一定的普适性, 我们同时将该方案放入无标度网络和小世界网络进行对比验证, 图 4 (c), (d) 分别表示了在无标度网络和小世界网络的高介数中心控制力节点和低接近中心控制力节点的抑制效果和病毒早期攻击传播速度. 可以看到在早期阶段无标度网络的感染规模和病毒传播速度都要明显高于小世界网络, 但是传播稳定后的感染数目则小世界网络稍微略大于无标度网络, 该现象可以解释成网络的异构化程度决定了网络传播特性, 在初始阶段由于小世界网络具有较多的强簇团小网络, 不利于疫情的快速传播, 但是到了疫情后期阶段, 网络的同质化程度越高反而越有利于疫情的长期传播.

4.2 分簇环境下重要节点对蠕虫病毒抑制和扩散的性能分析

文献 [16] 提出了簇系数对于节点获取新的邻居具有负面影响的理论, 证明了簇系数越大表明节点越容易在本簇团内部进行交流通信, 获取新节点的能力越差, 变相地阻碍了信息在网络中的传播. 因此本文采用随机游走理论将 1000 个节点的无标度网络分簇成聚类个数 (简记为 cluster) 分别等于 2 和 10 的簇团网络, 从而对分簇环境下重要节点对蠕虫病毒的抑制和传播扩散进行研究. 簇类系数 cluster = 2 和 cluster = 10 的网络形式如图 5 所示, 其中不同的节点颜色表示不同的簇团网络, 在程序中我们将每个节点都添加一个簇团 ID 属性, 节点在每次选择感染目标时优先选择与自己簇团 ID 相同的邻居节点. 与前面一样, 分别选择具有高连接度、高介数中心控制力、低接近中心控制力的节点作为重要节点来分析分簇环境中他们对蠕虫病毒防御和攻击的性能效果.



(a) (b)

图 5 (网刊彩色) 无标度网络分簇图 (a) 参数 cluster = 2 的 1000 个无标度网络节点分簇图; (b) 参数 cluster = 10 的 1000 个无标度网络节点分簇图

图 6 (a) 显示了 cluster = 2 时高介数中心控制力节点部署安全策略后病毒的感染效果同样取得了明显的抑制, 与图 4 (a) 不同的是在经过 $k = 18$ 中间传播跳数后, 感染的节点总数从 930 个下降到 906 个, 类似地如图 6 (b) 所示, 从攻击方角度而言选择低接近中心节点作为攻击目标时恶意程序的

传播速度也被减缓, 同样达到传播速度最大值的时间要比未分簇时来的迟缓. 图6 (c), (d) 显示了cluster = 10的实验结果, 与cluster = 2的结果相比较由于网络被分成大量的簇团子网, 蠕虫病毒的感染效果和传播速度都得到了很大程度的抑制, 高介数中心节点部署同样安全策略后经过$k = 18$的中间传播跳数后感染节点总数只有556个, 远未达到感染的饱和状态; 而同样选择低接近中心节点作为早期感染攻击的目标则需要经过$k = 15$的中间传播跳数后方能达到传播速度的最大化, 且如果选择随机节点和高介数中心节点作为早期攻击目标则经过$k = 18$的传播跳数后仍然没有达到传播速度的最大化; 因此可以证明网络分簇也能够对防御部署起到一定的优化作用.

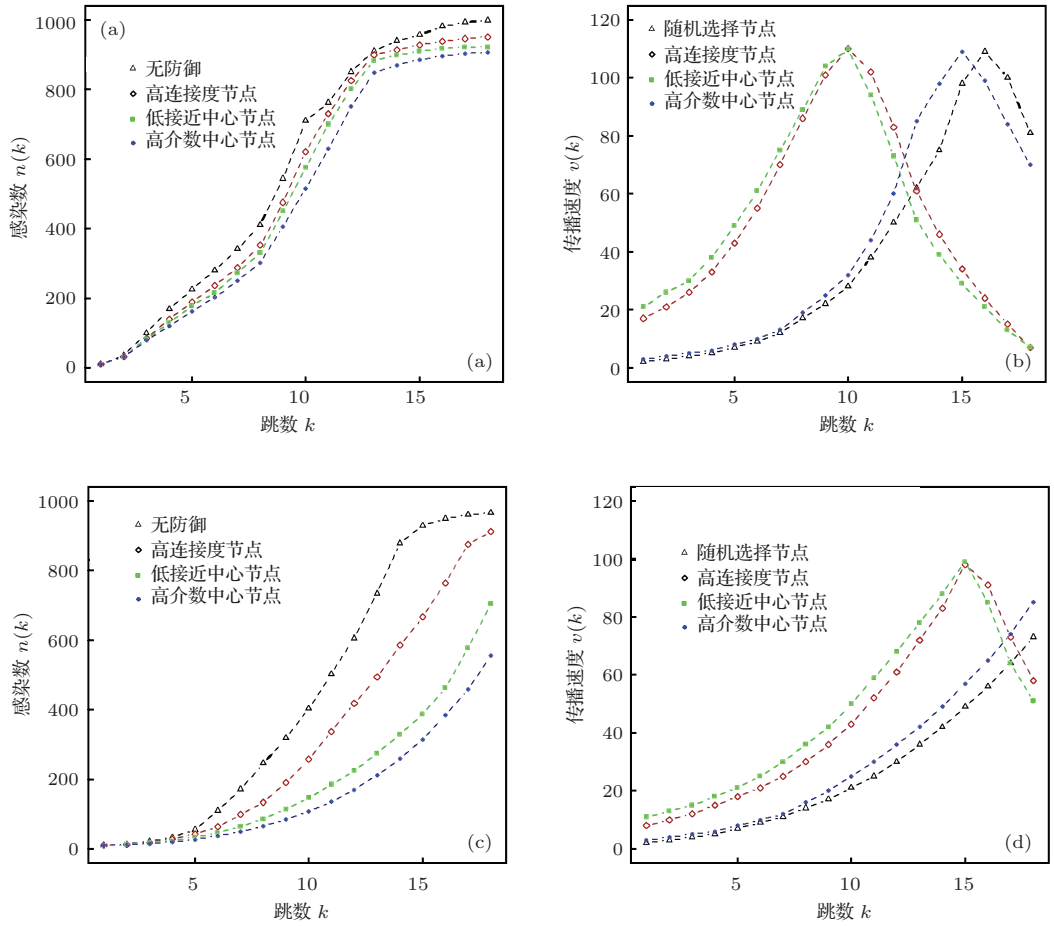


图6 (网刊彩色) 网络分簇环境下不同节点对恶意程序抑制和传播的效果分析 (a) 无标度网络中参数cluster = 2不同节点部署防御策略后的抑制效果; (b) 无标度网络中参数cluster = 2不同节点对病毒早期阶段传播速度的影响效果; (c) 无标度网络中参数cluster = 10不同节点部署防御策略后的抑制效果; (d) 无标度网络中参数cluster = 10不同节点对病毒早期阶段传播速度的影响效果

5 结 论

本文针对当前复杂网络中安全策略部署优化技术的不足, 提出了一种基于介数中心控制力和接近中心控制力模型的重要节点推荐部署技术. 该模型与以往研究工作强调高连接度节点是网络重要节点的理论不同, 引入了介数中心控制力和接近中心控制力概念, 实验结果证明在高介数中心节点和低接近中心节点上部署防御策略相对于传统的高连接度节点部署方法能够更有效地抑制和减缓病毒的感染疫情和传播速度; 同时本文也研究了网络的分簇行为对恶意程序感染和传播速度的影响, 证明了分簇对信息在网络中的传播具有抑制作用的观点. 当然需要指出的是, 本文提出的节点介数中心和接近中心的计算和网络环境有着非常紧密的联系, 当网络规模超过一定程度时候计算复杂度将会变得非常大, 这方面的研究也将会是未来该领域有待解决的一个研究课题.

参考文献

[1] 2014 Symantec Internet security threat report, Fossi M, Blackbird J http://www.symantec.com/content/en/us/enterprise/other_resources/b-istr_main_report_v19_21291018.en-us.pdf [2014-4-14]

[2] Zou C C, Towsley D, Gong W B 2007 *IEEE Trans. Dependable Sec. Comput.* **4** 105

[3] Fan X, Xiang Y 2010 *Future Gener. Comput.Syst.* **26** 1433

[4] Yan G H, Chen G L, Eidenbenz S, Li N 2011 *Proceedings of the 2011 ACM Symposium on Information, Computer and Communications Security* Hong Kong, China, March 22–24 2011 p196

[5] Xing Q B, Zhang Y B, Liang Z N 2011 *Chin. Phys. B* **20** 120201

[6] Song Y R, Jiang G P, Gong Y W 2012 *Chin. Phys. B* **21** 010205

[7] Lu Y L, Jiang G P, Song Y R 2012 *Chin. Phys. B* **21** 100207

[8] Maksim K, Lazaros K G, Shlomo H, Fredrik L, Lev M, Stanley H E, Hernan A M 2011 *Nature Physics* **6** 888

[9] Schneider C M, Moreira A A, Andrade J S, Havlin S, Herrmann H J 2011 *PNAS* **108** 3838

[10] Zhao D W, Li L X, Peng H P, Luo Q, Yang Y X 2014 *Phys. Lett. A* **10** 770

[11] Song Y R, Jiang G P 2010 *Acta Phys. Sin.* **59** 705 (in Chinese) [宋玉蓉, 蒋国平 2010 物理学报 **59** 705]

[12] Ebel H, Mielsch L, Bornholdt S 2002 *Phys. Rev. E* **66** 035103

[13] Mislove A, Marcon M, Gummadi K P, Druschel P, Bhattacharjee B 2007 *Proceedings of 2007 ACM/USENIX Internet Measurement Conference* California, USA, October 24–26 2007 p29

[14] Zou C C, Gong W B, Towsley D, Gao L 2005 *IEEEACM Trans. Networking* **13** 961

[15] Newman M E J 2005 *Social Networks* **27** 39

[16] Chen D B, Gao H, Lü L Y, Zhou T 2013 *PLoS ONE* **8** 77455

Recommendation of important nodes in deployment optimization model of defense strategy*

Yang Xiong[1)2)] Huang De-Cai[1)†] Zhang Zi-Ke[3)]

1) (*School of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023, China*)

2) (*School of Computer &Information Engineering, Changzhou Institute of Technology, Changzhou 213002, China*)

3) (*Alibaba Research Center for Complexity Sciences, Hangzhou Normal University, Hangzhou 311121, China*)

(Received 26 July 2014; revised manuscript received 10 October 2014)

Abstract

Current network security defense strategy focuses on deploying to high degree nodes where there are mainly two aspects of the problem: One is that the high-degree nodes are not the backbone nodes for the network communication in many occasions; another is that these nodes are not always the most effective ones for forwarding and propagation information. With the disadvantage of current network defense strategy deployment, this paper tends to improve the traditional diffusion model of malicious program propagation and measure the importance of network nodes by using intermediate hops, then the important node for recommended deployment technology based on betweenness center control and closeness center control model is put forward. Experimental results show that the nodes with high betweenness centrality and low closeness centrality as compared with the high degree nodes can more effectively quarantine the spreading of the worms whether in scale-free network or in small world network. Meanwhile, the clustering behavior of a network will also play a certain negative impact on the spread of malicious programs.

Keywords: high-degree nodes, betweenness centrality, closeness centrality, clustering

PACS: 05.10.–a, 02.30.Oz **DOI:** 10.7498/aps.64.050502

* Project supported by the Natural Science Foundation of Zhejiang pvovince, China (Grant Nos. LY12A05003, LQ13F030015), and the Natural Science Foundation of the Higher Education Institutions of Jiangsu Province, China (Grant No. 13KJD520001).

† Corresponding author. E-mail: hdc@zjut.edu.cn