

P2P 网络顽健性增强的方法

赵昊, 林伟, 刘胜利

(信息工程大学数学工程与先进计算国家重点实验室, 河南 郑州 450000)

摘要: 随着互联网的广泛应用, 网络通信管理架构和服务提供的稳定性愈发重要。构建一种基于邻居-邻居列表的 P2P 网络模型, 针对性地提出网络修复和修剪两种算法以提高网络结构的可靠性和顽健性。仿真实验表明, 在给定的威胁条件下, 提出的模型和算法可以有效提高 P2P 网络的自修复性。

关键词: P2P 网络; 邻居-邻居列表; 顽健性

中图分类号: TP311

文献标识码: A

doi: 10.11959/j.issn.2096-109x.2019020

Method for robust enhancement of P2P network

ZHAO Hao, LIN Wei, LIU Shengli

State Key Laboratory of Mathematical Engineering and Advanced Computing, Information Engineering University, Zhengzhou 450000, China

Abstract: With the widespread use of the Internet, the stability of network communication management architecture and service provision has become increasingly important. A P2P network model based on neighbor-neighbor lists was constructed, and two algorithms(network repairing and pruning) were proposed to improve the reliability and robustness of the network structure. Simulation experiments show that the proposed model and algorithm can effectively improve the self-healing of P2P networks under given threat conditions.

Key words: P2P network, neighbor-neighbor lists, robustness

1 引言

近年来, 网络技术发展迅速, 特别是进入“互联网+”^[1]时代以后, 日常生活的方方面面几乎都离不开网络。然而, 由于 TCP/IP 本身的不完善和

网络设备不可控等主客观条件所限, 网络服务的稳定性一直有待加强。互联网上的设备分布在各个地理位置, 通过特定的通信管理架构连接在一起, 保证其网络的连通性是确保网络服务稳定运行的前提条件。

收稿日期: 2018-11-10; 修回日期: 2018-12-20

通信作者: 赵昊, 754094629@qq.com

基金项目: 国家重点研发计划基金资助项目 (No.2016YFB0801505, No.2016YFB0801601)

Foundation Item: The National Key R&D Plan Program of China (No.2016YFB0801505, No.2016YFB0801601)

论文引用格式: 赵昊, 林伟, 刘胜利. P2P 网络顽健性增强的方法[J]. 网络与信息安全学报, 2019, 5(2): 88-94.

ZHAO H, LIN W, LIU S L. Method for robust enhancement of P2P network[J]. Chinese Journal of Network and Information Security, 2019, 5(2): 88-94.

中心结构的网络架构具有“单点失效”的固有脆弱性, 尽管 Domain-Flux^[2]和 Fast-fluxing^[3]等技术可以起到保护中心服务器的作用, 但并没有从根本上改变其集中式的本质。P2P 网络架构因此成为学术界和工业界的研究热点, 其分布式的拓扑结构以及对等节点的特性很好地规避了单一中心服务器的缺陷。

顽健性是评价网络性能的一个重要的指标。大量相关的论文从不同的方面对网络结构的顽健性进行了阐述和分析。因为网络中的节点分布在不同物理位置, 它们可能是普通机器也可能是服务器, 因此它们的在线时间存在不确定性, 有些节点可能频繁地加入或者退出网络, 而有些节点可能白天在线, 夜晚下线。无论是节点临时退出还是永久下线, 在网络管理者看来都是离线。在 P2P 网络中, 顽健性可等同于失去部分节点后, 最大限度地减小离线节点对整个网络的影响。

现有的 P2P 结构网络一般直接采用已知的网络协议, 这种方法具有实现相对简单的优点。然而, 提供稳定服务的 P2P 网络管理结构相比普通 P2P 网络在功能上有更高的要求。普通的 P2P 网络不需要全局信息, 如所有节点的数量, 而这对 P2P 管理网络来说十分重要。另外, P2P 管理网络的协议需要更加健壮, 因为所有网络节点可能随时被移除。基于以上观点, 本文提出一种改进的 P2P 网络拓扑结构, 具有自我修复的特性, 相对于普通 P2P 结构, 增强了自愈性和顽健性, 同时针对性地提出修复和修剪算法, 有效提高网络的连通性和防御能力。

2 P2P 网络概述

纯中心的网络结构虽具备实现简单、高效、协同性好等优势, 但其控制流程存在中心节点失效问题 (single of failure), 一旦攻击人员关闭中心服务器, 网络节点将无法继续获取命令, 管理

者也将失去对节点主机的控制权。为了提升网络的健壮性、对抗攻击人员的关停, 网络管理者使用非中心的 P2P (peer-to-peer) 模式作为其信道拓扑结构^[4], 如图 1 所示, 网络中每个节点主机既充当客户端又充当服务端, 通信过程不依赖公网可达服务器资源。不同于传统中心结构网络在域名或服务器被夺取控制权时表现出的脆弱, P2P 网络会建立一个相对分散的网络环境, 所有的节点主机都互相连接并且进行通信。虽然 P2P 网络数据发送延迟高于中心结构, 但不依赖脆弱中心节点^[5]的特性使其难以被整体劫持、测量和关闭。

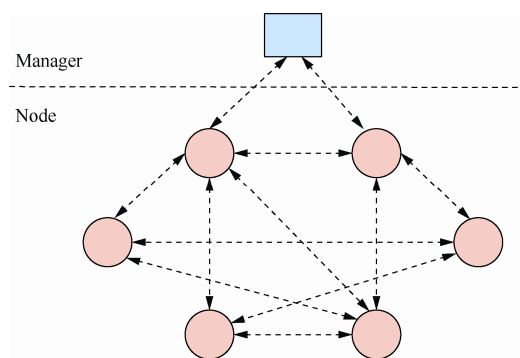


图1 P2P 网络结构

现有的 P2P 架构和协议虽然一定程度上缓解了中心架构的固有脆弱性, 但在设计之初并没有考虑到现在复杂多变的网络环境, 不能很好地抵抗当前严重的网络安全威胁。一般的 P2P 应用如文件共享、数据下载等, 虽然某些节点被移除, 对网络提供服务的影响不大, 然而一旦失效节点数量不断增加, 就会导致整个网络崩溃。另外, 某些对网络稳定性具有更高要求的应用, 如集群控制、通信管理等, 这些网络中的个别节点由于故障或者受到攻击而失效时, 可能会影响网络连通性从而导致网络管理者的控制力度不足, 无法实现预期的功能。因此, 提高 P2P 网络的顽健性, 使网络在失去部分节点后仍然能够实现较高的连通性并且尽量减少其他负面影响, 具有十分重要的研究价值和现实意义。

3 自修复 P2P 网络设计与实现

3.1 设计目标

P2P 网络通过其对等节点的结构,能够更好地提供网络通信、文件下载和数据传输等服务,提高传送速度和效率。P2P 网络最重要的特性即节点间互为客户端和服务端,因此普通的 P2P 网络在个别节点失效的情况下不影响其功能。然而,在需要提供更稳定持续服务的 P2P 网络中,节点失效造成的影响不容忽视。大规模的 P2P 管理网络中的节点在线率通常只有 10%左右^[6],因此每个在线节点都显得很重要。另外,直接与网络管理者相连的节点称为二级管理节点,在网络中起重要作用,直接接收网络管理者的命令并下发指令。

总体来说,影响网络健壮性的因素主要有以下 3 点。

1) 节点的自然关闭失效,如有的主机只在白天上线晚上会被切断电源。

2) 节点被人为攻击移除,网络中的主机很多处于无人维护状态,容易受到攻击。

3) 普通节点失效造成的影响只是局部的,一旦二级管理节点被移除,对整个网络的破坏是巨大的。

综上,普通 P2P 网络缺乏对网络服务提供健壮性的属性。本文的研究思路是通过连接方式的变化来提高 P2P 网络的顽健性。需要指出的是,无论是人为攻击还是自然关闭,本文只关注节点失效这一现象,不考虑节点被接管或污染等情况,将攻击威胁限定在节点被移除网络这一范围内。

3.2 邻居—邻居网络设计思想

P2P 网络通常采用的 bootstrap 机制主要有随机扫描方式和节点列表(peer-list)方式。基于随机扫描方式的 P2P 网络存在流量异常的固有脆弱点;采用基于 peer-list 交换方式进行通信,每台网络主机需要维护一份包含节点信息的列表,即

节点列表并定期随机挑选其中的部分节点进行通信,获取管理者指令和更新节点信息,这类 P2P 网络具有较好的灵活性、可扩展性、健壮性,在许多流行的 P2P 僵尸网络中使用^[7]。

图 2 是一个基于节点列表构建的 P2P 网络,共有 12 个网络节点。每个节点维护的节点列表信息如图 3 所示,如节点 1 存有节点 2、3、7 的地址,因此在 P2P 网络构建初始阶段就与这 3 个节点建立通信连接,接收管理者的命令和动态更新网络节点信息。

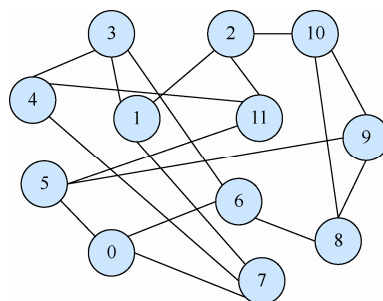


图 2 P2P 网络节点连接示意

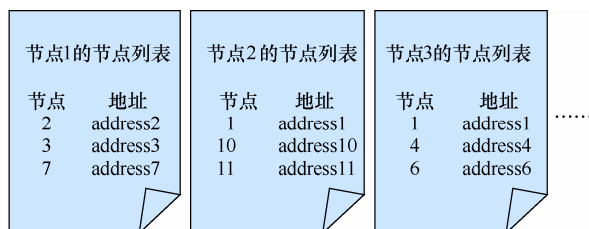


图 3 节点列表

实际运行中,网络节点存在被污染劫持而失效的可能性,而由于 P2P 网络中节点间的对等性,节点数量的损失很容易造成整个网络服务瘫痪。因此,如何在损失部分网络节点的情况下尽可能保证网络的有效性是值得研究的问题。一方面,需要使剩余节点建立通信连接,最大限度保证网络服务运行;另一方面,当节点数量减少后,主机节点间通信更为频繁,较容易被流量分析发现而被黑客攻击。

针对以上现实情况及理论分析,本文提出一种分布式动态自修复的网络拓扑结构。文献[8]研

究了 P2P 网络中的 NoN 贪婪路由,可减少路由长度,并且被证明是渐近最优的。本文讨论如何使用 NoN 的概念来创建一个自愈网络。

定义 1 邻居-邻居图: 考虑具有 n 个节点的图 G , 节点集合记为 V , 节点记为 $u_i (0 \leq i < n)$ 。 u_i 的邻节点集合记为 $N(u_i)$, 且 u_i 知道 $N(u_i)$ 包含的节点信息, 即每个节点均掌握其邻居的邻居节点身份及地址 (如 IP)。

仍以图 2 中的网络连接图为例, 图 3 中节点 1 的节点列表在邻居-邻居图中的形态如图 4 所示。节点 1 的邻居节点为节点 2、3 和 7, 其中节点 2 的邻居节点为 1、10 和 11, 节点 3 的邻居节点为 1、4 和 6……以此类推, 即邻居节点相应的邻居节点信息均存在与节点 1 的节点列表中, 这就是邻居-邻居图的设计思想, 是本文所提出的分布式自修复拓扑结构的构建基础。接下来重点讨论在邻居-邻居图模型上实现自修复拓扑结构的具体算法。

节点1的节点列表		
节点	地址	邻节点-地址
2	address 2	1-address 1, 10-address 0, 11-address 11
3	address 3	1-address 1, 4-address 4, 6-address 6
7	address 7	0-address 0, 1-address 1, 4-address 4

图 4 邻居-邻居节点列表示意

3.3 算法设计与描述

第 3.2 节指出提升 P2P 网络顽健性的两个要点, 下面就解决这两方面问题, 在邻居-邻居图模型下提出针对性的两个算法。

首先要考虑当有节点被移除时, 余下节点如何自发形成新的通信连接从而继续发挥网络服务的效能。在 P2P 网络中, 每个节点既是客户端也是服务端, 因此一旦失效, 尤其是关键节点的失效, 对网络的连通性会造成毁灭性打击。本文在邻居-邻居图的基础上提出网络修复算法, 利用邻居列表中邻居节点的信息, 使被移除节点的邻居节点间形成新的连接, “代替”被移除节点的中继

作用, 尽可能维持整个网络的连通性。修复的具体操作为: 当一个节点 u_i 被删除时, 判断其邻节点 u_j 和 u_k 是否可达, 若否, 则将 u_j 、 u_k 形成的边 (u_j, u_k) 加入现有边集合 E 。算法的流程描述如下。

算法 1 网络修复算法

- 1) for 每个被删除的节点 u_i
- 2) for u_i 的邻节点 u_j, u_k
- 3) if u_j 和 u_k 不可达
- 4) 生成边 (u_j, u_k) ;
- 5) 更新节点列表;
- 6) end for
- 7) end for
- 8) 完成修复过程

网络修复的具体过程如图 5 所示。节点 7 的邻居节点为 0、1、4, 当节点 7 被移除时, 节点 0、1、4 遍历自身的邻居列表判断互相之间是否连接, 然后形成新的边 (0、1)、(1、4) 和 (0、4)。之后, 当节点 6 被移除时, 重复相同的过程, 形成了新边 (0、3)、(3、8) 和 (8、0)。

接下来考虑另一方面的问题。对于图 G , 当节点 u_i 被删除时, 其每一个节点开始如上所述的修复过程。然而, 这会导致 u_i 邻节点的度数 (可以简单理解为节点的邻节点数量) 在 t 步之后显著增加, 如图 6 中的节点 0, 在经历节点 6 和 7 被移除的修复过程后, 其度数从 3 增加为 5。这会显著增加节点 0 被检测的可能性。因此, 确定一个节点度数的最大值 $d_{\max}$, 使节点 u_i 的每个邻节点删除节点列表中度数最低的节点 (即断开与其的通信连接), 直到该邻节点的度数范围符合条件。节点修剪算法描述如下。

算法 2 节点修剪算法

- 1) for 每个现有的节点 u_i
- 2) if 邻节点数量 $\geq d_{\max}$
- 3) 删除节点列表中邻节点最少的节点 u_j ;
- 4) 去除边 (u_i, u_j) ;

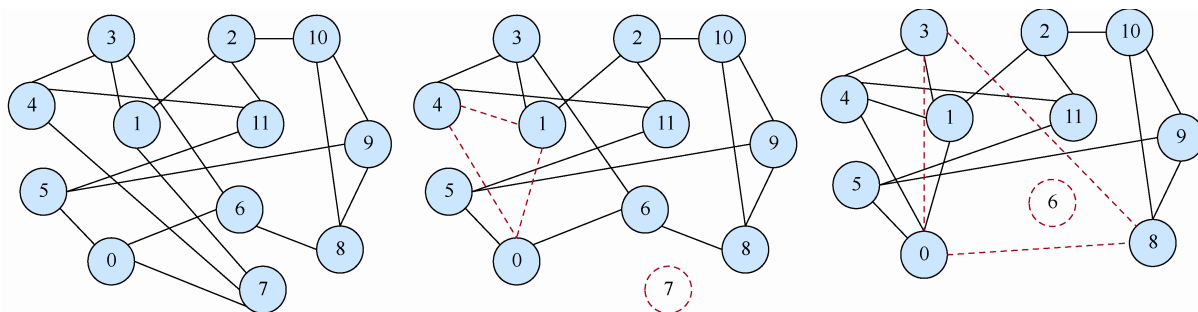


图5 修复过程

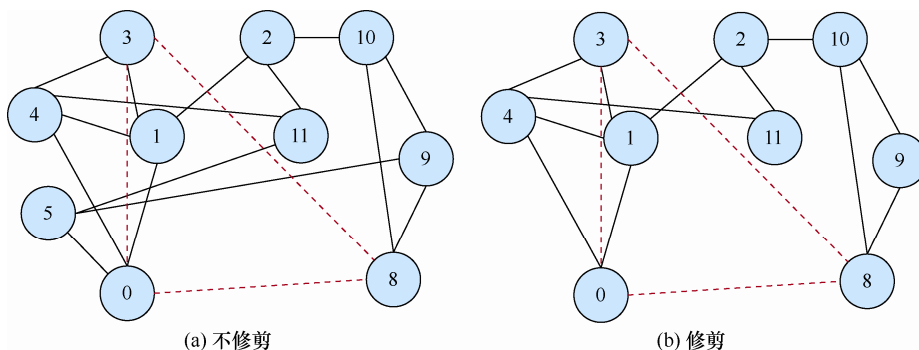


图6 修剪过程

- 5) 更新节点列表;
- 6) end for
- 7) 完成修剪过程

本文取 $d_{\max}=5$ 说明修剪算法的具体过程。如图6所示, 经过如上所述修复过程后, 节点0的度数为5, 触发了修剪过程的执行。此时, 节点0的邻节点1、3、4、5和8的度数分别为4、4、4、3和4。因此, 节点度数最低的5被删除。移除度数最低的节点是为了最大限度地保持剩余节点间的通信, 在一定程度上缓解节点数量减少的影响。

4 评估指标的确立

在对现有网络架构的顽健性进行研究时, 通常是通过模拟删除一定数量的节点后, 观察连接率的变化。连接率的影响越少, 说明网络中删除的节点对余下的节点影响不大, 因此顽健性越好。然而对于真实的网络环境, 通过删除一定的节点来研究其顽健性是不可能的。事实上, 分布式 P2P 网络本质上可看作复杂网络, 而对于复杂网络的

研究, 通常运用图论的相关知识^[9]。因此, 本文选取图论的属性和指标对 P2P 网络顽健性进行评估。然而, 图论中现有的属性如度数、最大距离等并不能直接作为 P2P 网络性能的指标。

为了对提出的邻居网络模型及两个算法的有效性进行测试, 评估其提高 P2P 网络结构顽健性的能力, 本文提出中心度和连接度两个属性作为评估指标。

定义2 中心度: 节点 u 到所有 $n-1$ 个其他节点之间的最短路径之和的倒数。

中心度主要反映网络中消息从节点 u 扩散传播至所有其他节点的速度, 因此是路径距离之和的倒数。考虑到距离总和取决于节点的数量, 所以计算时需要归一化处理, 计算公式如式(1)所示, 其中, n 是节点的数量, $d(u, v)$ 是节点 u 和 v 之间的最短路径。

$$C(u) = \frac{n-1}{\sum_{v \neq u} d(u, v)} \quad (1)$$

实验中, 单个节点的中心度反映不了整个网

网的情况,因此在测试时采取整个网络的平均中心度,由式(2)计算得到。

$$C = \frac{\sum C(u_i)}{n} \quad (2)$$

定义 3 连接度: 节点 u 的度数与最大可能度数的比值。

连接度表示网络中传播的消息流经节点 u 的可能性,实际上反映了网络通信流量的密集性。连接度计算公式如式(3)所示,其中, $d(u)$ 表示节点 u 的度数, k 表示网络节点最大度数。

$$D(u) = \frac{k}{d(u)} \quad (3)$$

在实验中,所有节点的平均连接度可作为测试指标,平均连接度可由式(4)计算得到。

$$D = \frac{k n}{\sum d(u_i)} \quad (4)$$

5 实验测试与分析

虽然理论分析和验证更为严谨,但难度更大。因此本文采用仿真实验来评估基于邻居列表的 P2P 网络特性。实验在 500 个节点的 k -正则图($k=5, 10, 15$) 中模拟节点删除过程,最多有 30% (150) 个节点被删除。

图 7 显示了修剪(如图 7(a) 所示)和不修剪(如图 7(b)所示)时网络中节点的平均中心度随删除节点数量的变化。正如图 7 所反映,节点中心度是稳定的,并且即使节点删除后它也不会减少。另外,未修剪时,由于剩余节点间修复过程的作用未被抵消,因此平均中心度显著上升。实际上,中心度反映消息传播速度,只要满足基本要求,大于或等于初始速率即可。实验结果表明,本文提出的网络模型对网络中心度无负面影响。

图 8 反映了修剪(如图 8(a)所示)和不修剪(如图 8(b)所示)时网络中节点的平均连接度随删除节点数量的变化。显而易见,若在修复过程

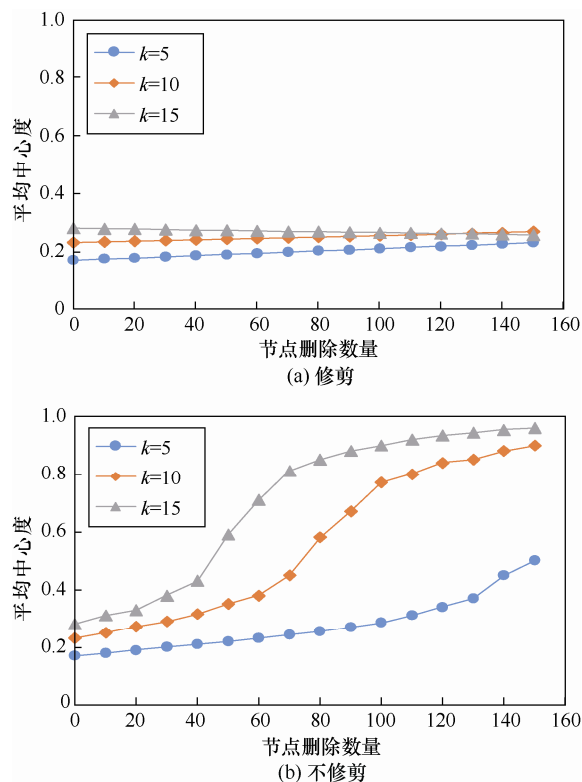


图 7 中心度变化

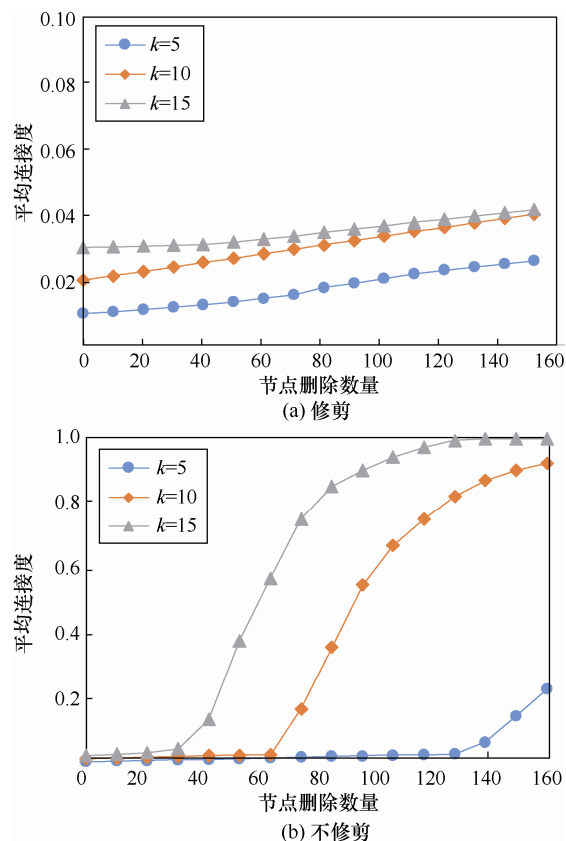


图 8 连接度变化

后不增加修剪过程,节点的连接度会大幅增加,这说明整个网络的消息传播率显著提升。然而,网络的通信频率大幅增加会导致节点间的通信流量明显增多,造成的后果是容易被检测和针对,使网络暴露,最终被攻击而损毁。

6 结束语

互联网已成为人们工作和生活中不可或缺的一部分。如何提高网络服务的稳定性已成为亟待解决的问题。本文研究了网络拓扑的改进,提出了一种分布式、自愈合的P2P网络体系结构。通过网络修复和修剪算法,增强了网络的健壮性和顽健性,可以有效抵抗节点删除的影响。下一步是将此模型应用于真实环境P2P网络服务,以验证其在实际条件下的可靠性。

参考文献:

- [1] 周锡冰. 互联网+服务[M]. 北京: 人民邮电出版社, 2016.
ZHOU X B. Internet + Service[M]. Beijing: Posts and Telecommunications Press, 2016.
- [2] ALMOMANI A. Fast-flux hunter: a system for filtering online fast-flux botnet[J]. Neural Computing & Applications, 2016: 1-11.
- [3] SOLTANAGHAEI E, KHARRAZI M. Detection of fast-flux botnets through DNS traffic analysis[J]. Scientia Iranica, 2015, 22(6).
- [4] GRIZZARD J B, SHARMA V, NUNNERY C, et al. Peer-to-peer botnets: overview and case study[C]//USENIX Workshop on Hot

Topics in Understanding Botnets (HotBots'07). 2007:1.

- [5] FEILY M, SHAHRESTANI A, RAMADASS S. A survey of botnet and botnet detection[C]// International Conference on Emerging Security Information. 2009:268-273.
- [6] LI H, HU G, YANG Y. Research on P2P botnet network behaviors and modeling[C]//International Conference on Information Computing and Applications. 2012:82-89.
- [7] YIN J, CUI X, LI K. A reputation-based resilient and recoverable P2P botnet[C]//IEEE Second International Conference on Data Science in Cyberspace. 2017:275-282.
- [8] MANKU G S, NAOR M, WIEDER U. Know thy neighbor's neighbor: the power of lookahead in randomized P2P networks[C]//The 36th Annual ACM Symposium on Theory of Computing. 2004: 54-63.
- [9] MACARTHUR B D, SÁNCHEZ-GARCÍA R J, ANDERSON J W. Symmetry in complex networks[J]. Discrete Applied Mathematics, 2008, 156(18):3525-3531.

[作者简介]



赵昊(1993-), 男, 浙江诸暨人, 信息工程大学硕士生, 主要研究方向为网络与信息安全、大数据分析。

林伟(1986-), 男, 湖南常德人, 博士, 信息工程大学讲师, 主要研究方向为软件保护与分析、网络安全。

刘胜利(1973-), 男, 河南郑州人, 博士, 信息工程大学教授, 主要研究方向为机器学习、网络安全。