P2P网络数据污染综述

王 勇[1] 云晓春[2] 秦志光[1] 郭 莉[2] 程红蓉[1]
(电子科技大学计算机科学与工程学院 成都 610054)[1] (中科院计算技术研究所 北京 100080)[2]

摘 要 对等网络应用(Peer-to-Peer networking applications,P2P)相关的安全威胁已经成为广受关注的网络安全课题。P2P网络共享文件内容的数据污染给P2P网络安全带来新的难题:例如,如何快速定位发现污染数据,分析污染数据特征模型,寻求高效低代价的数据污染治理策略等还有待进一步深入分析研究。针对P2P网络数据污染,从污染数据监测、数据污染特征模型分析以及数据污染治理策略等三方面,阐述了当前该邻域的主要研究动态,分析了数据污染相关研究的关键问题,最后指明了该邻域未来可能的发展方向。

关键词 对等网络,数据污染,网络测量,网络安全
中图法分类号 TP939 **文献标识码** A

Survey on P2P Network Pollutions

WANG Yong[1] YUN Xiao-chun[2] QIN Zhi-guang[1] GUO Li[2] CHENG Hong-rong[1]
(Dept. of Computer Science and Engineering, University of Electronic and Science Technology of China, Chengdu 610054, China)[1]
(Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080, China)[2]

Abstract In recent years, the security issues of Peer-to-Peer networking applications(P2P) are becoming the hot spots of network security research. The pollution and poisoning problems of shared content in P2P network present a new challenge for P2P networking security: such as how to find and locate the polluted data, how to modeling the pollution and poisoning process, and how to prevent the spreads of polluted contents efficiently etc. In this paper, the up-to-date researching results were discussed in detail from the three main angles as the polluted data measurements, the modeling processes, and the pollution constrain strategies; some key issues related to P2P network pollutions were analyzed. Finally, the paper pointed out the future possible researching trends on P2P network pollutions.

Keywords P2P network, Pollution and poisoning, Network measurements, Network security

1 引言

随着对等网络应用在Internet网络上的广泛流行,与之相关的诸多课题受到研究者和领域专家的密切关注。对等网络应用程序,通过充分利用网络中客户端的带宽、存储空间和计算能力,为用户提供可扩展的、高效的信息访问机制。作为一类大规模、自组织、高度动态的复杂网络系统,对等网络在Internet上的大量部署和应用,严重威胁着Internet网络的信息安全[1]。首先,P2P网络改变了原有的Internet流量模型,造成网络流量的极度拥塞,影响了ISP的服务质量;其次,大量利用P2P网络传播的病毒、木马已经成为Internet上的一个巨大安全隐患;最后,对等网络中的共享资源,往往缺乏有效的监管机制,大量的数字媒体被非法传播,或受到恶意篡改,给知识产权拥有者造成巨大的经济损失。近年来,音乐媒体公司纷纷发起对P2P文件共享系统的声讨,寻求各种P2P网络系统的封杀方案,包括利用法律起诉P2P网络服务公司,追踪P2P网络下载用户,甚至雇用专门公司使用技术手段,对P2P网络发起大规模恶意攻击等等。

P2P网络中的数据污染已经成为一个重要的安全问题,主要安全威胁表现在:(1)P2P网络中存在的污染信息,严重降低了P2P网络系统的性能,受污染的文件可能携带病毒、木马等恶意程序,威胁P2P网络终端用户的安全;(2)由恶意攻击者发起的数据污染攻击,可能造成P2P网络系统中大量节点失效,进而使整个P2P网络崩溃,特别地,数据污染攻击已经成为基于离散哈希表(DHT)的结构化对等网络系统最大的安全威胁之一;(3)由数据污染引发的拒绝服务攻击,可以湮没P2P网络中的共享资源,增加Internet网络的流量负担,从而降低ISP对Internet用户的服务质量。

近两年,P2P网络中的数据污染问题,受到P2P网络系统开发商、Internet服务提供商(ISP),以及信息安全研究人员等多方密切关注,对数据污染攻击模型化研究的需求来源于以下几个方面:(1)通过解决P2P网络中大量存在的数据污染问题,可以增强P2P网络应用的性能和安全性,为网络用户提供高效的共享资源定位和分发服务;(2)掌握P2P网络中数据污染攻击的特征、传播方式、影响范围等知识,能够为ISP进行有效的流量规划和用户服务质量管理提供依据,提高Internet网络的使用效率;(3)从国家网络信息安全角度考虑,管理监管部门需要了解P2P网络中数据污染问题对信息

到稿日期:2010-04-26 返修日期:2010-10-09 本文受国家242信息安全计划(2008A16),国家863计划(2009AA01Z422)资助。
王 勇(1976—),男,博士,副教授,主要研究方向为网络安全、P2P网络技术,E-mail:cla@uestc.edu.cn。

安全的危害性，需要在线管控 P2P 网络行为，促进 P2P 网络应用协议的标准化进程，对数据污染攻击的深入分析是实现这些目标的前提。

P2P 网络数据污染攻击的模型化研究是一项复杂的工作，设计到网络测量、图论、算法设计、统计学、数据挖掘、可视化以及数学建模等多个研究领域。相应的研究内容可以归纳为 3 个问题：(1)如何有效地识别数据污染，准确而完整地获取 P2P 网络应用系统中数据污染状态数据；(2)如何对 P2P 网络中数据污染扩散特征进行描述；(3)如何有效地抑制 P2P 网络中的数据污染，防范数据污染攻击。这 3 个问题分别对应于 3 个研究方向，即：P2P 网络数据污染测量、P2P 网络数据污染特征分析、P2P 网络数据污染治理。其中，P2P 网络测量是数据污染特征分析的基础；数据污染特征分析是数据污染模型化的核心；P2P 网络数据污染防御是模型化的应用目的。本文将对这 3 个研究方向分别进行论述。

2 P2P 网络数据污染监测

P2P 网络数据污染测量分类方法有多种：根据测量方法可以分为主动测量和被动测量；根据测量系统自身的部署情况，可以分为单点测量和分布式多点测量；根据对污染数据的判别算法，可以分为基于消息统计的污染测量、基于比较分析的污染测量、基于智能学习的污染测量等。本节根据测量对象的不同，可以分为文件型数据污染测量和流媒体数据污染测量。

文件型数据污染测量是目前最为关注的数据污染测量问题。2005 年，KaZaA 网络中的蓄意污染攻击已经造成系统中 50%～80%的文件受到污染[2]；eDonkey 网络中有 50%左右的热门文件受到索引毒害攻击(index poisoning attack)[3,4]；Seungwon 等人[5]建立了 eDonkey 网络爬行器-Krawler，测量分析了 P2P 病毒在 eDonkey 网络中的流行情况；UC Berkeley 的 Christin 等人[6]采用主动/被动相结合的方式测量了 4 类主流 P2P 文件网络(eDonkey，Overnet，FastTrack 和 Gnutella)中的数据污染，从共享文件的可用性、稳定性、下载时间等几个方面分析了这 4 类 P2P 网络的污染态势，此外，还分析了网络拓扑结构与数据污染扩散之间的内在联系。Dhungel 等人[7,8]通过构造 BT 网络 utorrent 客户端爬行器(主动测量)和分析 BT 网络 Tracker 服务器日志信息(被动测量)的方法，测量分析了 Bittorrent 网络中 leechers 节点遭受数据污染攻击的情况，分别分析了 BT 网络数据污染的安全现状。

随着基于 P2P 技术的流媒体应用日益流行，P2P 流媒体网络中的数据污染状态也很不乐观。文献[9]实现了测量 PPLive 网络的爬行器(Crawler)，通过构造 PPLive 客户端程序参与流媒体业务，来获取网络中邻居节点的信息和消息数据。测量结果显示：PPLive 网络同样在遭受数据污染攻击，并且与网络中节点的物理位置、视频内容、在线时间、社会新闻事件等因素有密切的关系。

P2P 网络的大规模、自组织、高度动态等特性使数据污染测量变得非常困难。目前，对数据污染的测量仍处于观测、总结特性的阶段，还未形成完整的测量框架体系，此外，也未形成完整的测量数据可靠性、完整性、准确性的评价指标和方法。

3 数据污染模型

近年来，P2P 网络中数据污染问题越来越受到关注，相关的模型化分析主要关注数据污染的特征、扩散过程、敏感性、危害性等方面的属性。一般地，可以从文件版本污染攻击和网络毒害攻击两个方面建模来描述数据污染的扩散过程，相应的数据污染模型可以分为两类：一类是基于免疫学原理的经典模型，包括文献[10-12]中建立的模型；另一类通过描述节点状态的变迁来展现数据污染在 P2P 网络中的扩散过程，包括 Kumar 等人在文献[13]提出的流模型(fluid model)及 Shi 等人在文献[14]中建立的状态模型等。

3.1 基于免疫学原理的经典模型

对于第一类经典模型来说，是对 SIR 传染病模型的应用和拓展，将 P2P 网络中的节点分为 3 种，即：易感节点(Susceptible peers)、感染节点(Infected peers)和免疫节点(Recovered peers)。这 3 种节点的相互转换关系如图 1 所示。

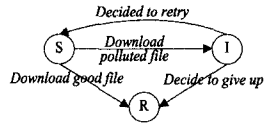

图 1 P2P 网络中 3 类节点的转换关系

由于 P2P 网络中的文件受到污染攻击，节点可能下载到原版文件和受污染的文件。当节点下载了受污染的文件版本时，它即加入感染节点群体；同时，节点通过再次尝试或放弃下载等方式，可以加入免疫节点群体。

设网络中对共享文件 i 感兴趣的节点数量为 M_i，初始时原版和污染版的文件数量分别为 $N_g(0)$ 和 $N_b(0)$，在时刻 t，网络中 3 类节点的数量分别为 $S(t)$，$I(t)$ 和 $R(t)$，通常：$S(0)=M_i$，$I(0)=R(0)=0$。则可以建立如下微分方程组：

$$\begin{cases} p_b(t)=\dfrac{N_b(t)}{N_b(t)+N_g(t)} \\ \dfrac{dS(t)}{dt}=-\lambda_s S(t)+\lambda_r I(t) \\ \dfrac{dI(t)}{dt}=p_b(t)\lambda_s I(t)-(\lambda_r+\lambda_x)I(t) \\ \dfrac{dR(t)}{dt}=(1-p_b(t))\lambda_s S(t)+\lambda_x I(t) \\ \dfrac{dN_b(t)}{dt}=\lambda_s p_b(t)p_{sb}S(t)-(\lambda_r+\lambda_x)p_{sb}p_{ib}I(t) \\ \dfrac{dN_g(t)}{dt}=\lambda_s(1-p_b(t))p_{sb}S(t) \end{cases} \tag{1}$$

基于免疫学原理的经典模型，通过选择合适的微分方程组因数，能够较好地描述版本污染和毒害攻击在 P2P 网络中的扩散过程。然而，P2P 网络作为复杂网络系统的一个实例，其网络拓扑结构和传播动力学特性是影响数据污染扩散的重要因素，该类经典模型认为 P2P 网络自身是个全连接拓扑图，忽略了拓扑特性对污染扩散的影响；另一方面，P2P 网络的动态性使其在网络传播动力特性方面与 Internet 有很大的区别，经典模型在这一方面的描述能力受到较大的限制。此外，经典模型不能反映 P2P 网络客户端节点共享行为异构的特性，降低了经典模型描述的准确性。

对该经典模型进一步拓展，以描述 P2P 网络中毒害攻击的传播过程，其中一个典型的毒害攻击扩散离散模型表示如下：

设共享目标文件的大小为 f_{size}，被 P2P 网络系统分割为 m 个 pieces(典型地，$m=f_{size}/256$)；

设 λ 为共享文件在网络中被下载的概率，即：文件在 P2P 网络中的流行程度；η 为节点共享文件一段时间后节点下线或停止共享服务的概率；

令 $D_0, D_1, \cdots, D_{m-1}$ 分别表示节点已经成功下载了第 i 个 piece，特别地，D_0 表示节点完成下载的初始化，开始进入正式下载状态，D_{m-1} 表示节点已经成功下载完所有 m 个分片，完成了文件下载过程；

令 S 和 I 分别表示节点处于共享状态(提供下载服务)和空闲状态(不共享下载的分片或者离开 P2P 网络)；

为简化模型，假设节点下载完全部分片后才以概率 p_{share} μ 进入共享状态。节点状态的迁移模型如图 2 所示。

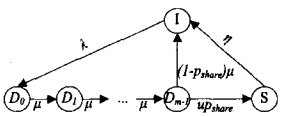

图 2　毒害攻击下节点的状态迁移模型

当节点下载完毕一个分片后，客户端使用带宽优先和轮叫算法选择节点下载新的分片，设：P2P 网络中节点的平均上传带宽和下载带宽分别为 r_u，r_d，则节点获取下一分片的概率可以表示为：$\mu(t)=\frac{f_{size}}{m}\min\left\{\frac{S(t)r_u}{\sum_{i=0}^{m-1}D_i}, r_d\right\}$。时刻 t 各状态节点数量变化可以用如下离散微分方程组表示：

$$\begin{cases}\dfrac{dI(t)}{d(t)}=(1-p_{share})\mu(t)D_{m-1}(t)-\lambda(t)I(t)+\eta(t)S(t)\\ \dfrac{dD_0(t)}{d(t)}=\lambda(t)I(t)-\mu(t)D_0(t)\\ \dfrac{dD_i(t)}{d(t)}=\mu(t)(D_{i-1}(t)-D_i(t)) \quad \forall 1\leqslant i\leqslant m-1\\ \dfrac{dS(t)}{d(t)}=p_{share}\mu(t)D_{m-1}(t)-\eta(t)S(t)\end{cases} \tag{2}$$

式中，网络中的总节点数 $N=I+\sum_{i=1}^{m}D_i+S$ 不变，并且 $N=I(0)+S(0)$，$D_i(0)=0$，$i=0,1,\cdots,m-1$。

目前，该模型还没有考虑 P2P 网络中节点的动态特性，也没有考虑节点下载完成一个分片后立即共享的情况，在进一步的建模研究中，将在此模型基础上进行细化，融入 P2P 网络的结构特征，形成 P2P 网络毒害攻击的离散模型。

3.2　基于状态迁移的模型

第二类状态模型建模针对 P2P 网络应用中节点的行为迁移特性，研究污染版本和毒害攻击的扩散过程。Kumar 等人建立了对等网络中数据污染扩散的流模型(Fluid Modeling)，分析了流模型下客户端行为(包括主动放弃下载、黑名单过滤等)对 P2P 网络中数据污染扩散的影响。结果显示：目前单纯的黑名单、信誉系统等策略不能有效抑制污染的扩散。Kumar 的流模型为 P2P 网络数据污染扩散建模提供了基本的理论思路。模型考虑了“自由下载(free loading)”和“用户中途放弃”等两种用户行为对污染信息扩散的影响，此外，还分析了“非线性版本偏好”(non-linear bias toward popular versions)、黑名单等防治策略对数据污染扩散的影响。该模型描述如下：

设 $x(t)$，$y(t)$ 分别表示时刻 t 网络中原版文件(good copy)和“盗版”文件(polluted copy)的总数，它们均为连续时间函数，则此时用户选择下载“盗版”文件的概率 $p(t):=\frac{y(t)+N}{x(t)+y(t)+N}$，其中 N 是攻击者最初注入网络的盗版文件数量。

原版文件在网络中的数量 $x(t)$ 变化有两种可能：(1)没有该共享文件的用户选择下载了原版文件；(2)下载了“盗版”文件的用户重新选择下载了原版文件。因此，$x(t)$ 的流方程可以表示为：

$$\dot{x}(t)=[M-x(t)-y(t)]\mu(1-p(t))+y(t)\mu(1-p(t)) \tag{3}$$

类似地，“盗版”文件在网络中的数量 $y(t)$ 变化有两种可能：(1)没有该共享文件的用户选择下载了盗版文件；(2)下载了“盗版”文件的用户重新选择下载了原版文件。因此，$y(t)$ 的流方程可以表示为：

$$\dot{y}(t)=[M-x(t)-y(t)]\mu p(t)-y(t)\mu(1-p(t)) \tag{4}$$

式中，M 为网络中最初没有该共享文件副本的节点总数，μ 为节点校对文件是否原版的时间指数分布参数。

通过式(3)和式(4)，可以得到 $x(t)$ 和 $y(t)$ 的解为：

$$x(t)=\frac{c_2M(e^{\mu t}-\frac{c_1}{M+N})^{\frac{M}{M+N}}}{1+c_2M(e^{\mu t}-\frac{c_1}{M+N})^{\frac{M}{M+N}}} \tag{5}$$

$$y(t)=M-c_1e^{-ut}-x(t) \tag{6}$$

式中，$c_1=M-x(0)-y(0)$，

$c_2=\frac{x(0)}{M-x(0)}\left(\frac{N+x(0)+y(0)}{M+N}\right)^{-\frac{M}{M+N}}$。

图 3 描述了流模型中文件版本的状态迁移过程。

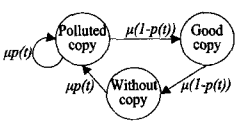

图 3　污染攻击下的状态迁移模型

该模型是 P2P 网络中文件污染扩散的最基本形态，通过对模型参数的调节和加入 BT 网络的动态特征、网络结构特征，可以进一步研究污染文件对 BT 网络的影响，发现网络中的脆弱组件。

4　数据污染治理技术

研究 P2P 网络数据污染测量技术和模型化方法的最终目的是寻求 P2P 网络环境下高效的数据污染治理技术。数据污染的治理包含两个层次，即：抑制污染信息扩散和清除污染信息。目前，还未见清除 P2P 网络环境下污染信息的系统方法，在数据污染抑制方面，采用的方法主要包括信任机制、数字签名技术和信息冗余等。

Kevin 等人[15]采用投票策略的信任机制防御 P2P 网络中虚假信息等污染的扩散；Cristiano 等人[16,17]提出了一个高效的分布式信任系统-Scrubber，用以抵御 P2P 网络中基于信息诱骗和身份修改的污染攻击。Fabricio 等人[18]研究了用户激励机制和用户行为对治理 P2P 网络数据污染的影响，其中，用户行为包括：(a) 用户主动清除行为；(b) Hash 伪造行为；(c)下载源数量限制行为等。模拟分析显示：针对不同的数据污染，需要采取对应的客户端激励机制才能减缓污染信息的扩散。Lou 等人[3]提出了可靠索引交换协议(Reliable

Index Exchange Protocol,RIEP),用以防止 P2P 文件共享应用中数据污染等造成的 DDoS 攻击,分析了 RIEP 的有效性和实现时的局限性。文献[19]提出将磁盘阵列中的 RAID 技术应用到 P2P 文件共享系统中,以降低数据污染对 P2P 网络系统的危害。在国内,对等网络的数据污染和治理方面的研究还处于初始阶段,理论和经验数据等方面的研究和分析还不深入。中科大的韦冬等人[20],利用向量空间相似度赋予投票权重,使用 horting 图方法解决数据稀疏性问题,采用自适应的信誉阈值判断文件可信性。

目前,P2P 网络中数据污染的治理还处于理论模型研究起步阶段,上述策略和方法还不能有效解决具有高度动态、自组织、大规模等特征的 P2P 网络中的数据污染问题。

5 难点问题和研究思路

P2P 网络数据污染研究目前面临的主要困难来自于 3 个方面:首先是 P2P 网络结构和行为复杂的特性;其次是数据污染自身的复杂性;最后是缺乏系统研究的方法论。

P2P 网络结构和行为复杂的特性带来的问题主要集中在数据污染测量方面,具体表现在:P2P 网络拓扑的动态变化影响测量的准确性;其庞大的规模影响了数据污染测量数据的完整性;节点异构和用户行为不确定增加了数据污染测量的技术难度。因此,客观地说,目前的测量技术无法获得完整、准确的 P2P 网络数据污染数据,现有的 P2P 网络数据污染模型都存在一定的局限性。由此引发的问题包括:P2P 网络中污染信息扩散与网络自身特点之间的内在联系是什么?回答这个问题不仅需要结合 P2P 网络流量、拓扑、可用性等多层次的测量分析,还需要引入用户行为特征、Internet 网络行为特征以及信息内容特征等多方面因素。Christin 等人[6]根据查询返回数量、查询响应时间、内容稳定性、内容重复度、下载完成时间等量化指标对 Gnutella, eDonkey, Overnet, FastTrack 这 4 个主流 P2P 网络的内容特征进行了测量,分析了各种污染策略对 P2P 网络内容可用性的影响。Shi 等人[14]通过分析 P2P 网络结构层次特征,结合数据污染策略的特点,面向节点行为进行建模分析,建立了 P2P 网络污染统一模型,进而分析了不同数据污染策略对 P2P 网络性能的影响。本文认为,结合上述特征,进一步研究 P2P 网络中数据污染扩散背后的驱动机制,从复杂网络系统的 HOT(Highly Optimized Tolerance)理论中寻求答案,是未来 P2P 网络数据污染模型化研究的一个重要方向。

就数据污染自身来说,信息的多样性使得智能识别污染非常困难;庞大的信息内容和信息的变化增加了污染信息识别的效率和准确性;而互联网络相关的版权、隐私保护等法律、社会因素,影响了数据污染治理的部署和实施。在污染信息智能识别技术方面,研究者提出了比较法、特征提取、节点消息校验等识别算法,但是它们在效率和准确性上都还不能满足 P2P 网络大规模和动态性的需求。研究在 P2P 网络环境下识别污染信息、判定信息内容的可信性是亟待解决的公开问题之一。

随着研究的不断深入,特别是对 P2P 网络自身特征的深入认识,人们发现 P2P 网络属于具有"新陈代谢"功能的自组织复杂系统,它与 Internet 网络、万维网、细胞网络、社会关系网、科学文献引用网络等具有相同的内在特征——自相似特性,同时认识到对 P2P 网络特性、网络中信息内容特性以及数据污染的特性等方面缺乏了解不只是一个计算机科学问题,而是源于缺少一个对复杂网络系统进行特征化分析的科学框架。因此,建立 P2P 网络特征模型、研究 P2P 网络中污染信息特征模型要依赖科学界在上述邻域的共同努力,对 P2P 网络模型、数据污染模型的研究也会促进这一科学框架的形成。

综上所述,P2P 网络数据污染的研究路线主要包括以下 3 个方面:(1)通过分析 P2P 应用系统的网络结构、内容分布、以及用户行为等特征,建立新的 P2P 网络数据污染测量技术,提高测量数据的完整性、准确性以及污染信息的识别效率,以求从新数据中发现数据污染新特征;(2)将数据挖掘、复杂网络以及图论结合起来,研究 P2P 网络数据污染/毒害攻击模型、数据污染扩散模型,建立系统的 P2P 网络数据污染特征分析方法;(3)借鉴复杂网络的研究成果,利用信任模型、密码学等方法,研究 P2P 网络环境下数据污染的抑制技术,建立多层次的 P2P 网络数据污染评价方法和指标体系。

参 考 文 献

[1] CNCERT/CC. 2007 年上网络安全工作报告 [R]. http://www.cert.org.cn/UserFiles/File/CNCERTCC2007AnnualReport_Chinese.pdf,2008-04-08

[2] Liang J, Kumar R, Xi Y, et al. Pollutions in P 2 P File Sharing Systems [A]// IEEE Infocom 2005 [C]. 2005:3

[3] Lou X, Hwang K. Prevention of Index-Poisoning DDoS Attacks in Peer-to-Peer File-Sharing Networks [R]. USC Internet and Grid Computing Lab(TR-2006-5). 2006,11

[4] Liang J, Naoumov N, Ross K W. The Index Poisoning Attack in P2P File-Sharing Systems [A]// Proceedings of IEEE Infocom 2006[C]. Barcelona, Spain. USA: IEEE Press, 2006,4

[5] Seungwon S, Jaeyeon J, Hari B. Malware prevalence in the KaZaA file-sharing network [A]// Proceedings of the 6th ACM SIGCOMM on Internet measurement [C]. Rio de Janeriro, Brazil, ACM Press, 2006

[6] Christin N, Weigend A S, Chuang J. Content Availability, Pollution and Poisoning in File Sharing Peer-to-Peer Networks [A]// Proceedings of the 6th ACM conference on Electronic commerce Vancouver [C]. BC, Canada: ACM Press, 2005:68-77

[7] Dhungel P, Wu D, Hei X, et al. Is BitTorrent Unstoppable? [A]// IPTPS 2008[C]. Tampa Bay, Florida, Feb. 2008

[8] Dhungel P, Wu D, Schonhorst B, et al. A Measurement Study of Attacks on BitTorrent Leechers [A]// IPTPS 2008[C]. Tampa Bay, Florida, Feb. 2008

[9] Dhungel P, Hei X, Ross K W, et al. The Pollution Attack in P2P Live Video Streaming: Measurement Results and Defenses [A]// Sigcomm P2P-TV Workshop[C]. Kyoto, 2007

[10] Thommes R W, Coates M J. Epidemiological Modelling of Peer-to-Peer Viruses and Pollution [A]// Proc. of IEEE Infocom[C]. Barcelona, Spain, 2006,4

[11] Dumitriu S, Knightly E, Kuzmanovic A, et al. Denial-of-service resillience in peer-to-peer file sharing systems [A]// SIGMETRICS'05 [C]. Banff, Alberta, Canada, 2005,6

[12] Liang J, Naoumov N, Ross K W. Efficient Blacklisting and Pollution-Level Estimation in P2P File-Sharing Systems [A]// AINTEC 2005[C]. 2005:1-21

(下转第 23 页)

4.9 电力市场动荡性研究问题

电力市场是电力企业的重要输出源，它通过市场化运作给电力企业带来竞争压力。

2008年，Sahraei-Ardakani等[9]采用IWO算法求解了电力市场动荡性问题，通过大量计算表明IWO算法能够在具有纳什均衡点的问题中求得纳什均衡解。

5 IWO算法的可扩展性探讨

由于IWO算法是一种数值优化算法，适合求解具有连续变量的函数优化问题。如果仅限于此的话，IWO算法的应用和推广将变得异常困难。如何扩展IWO算法，使其能够求解混合整数规划问题和带有二进制变量的函数优化问题，是非常值得考虑的重要方面。从算法的执行过程来看，其他部分都是框架性的，只有空间分布环节需要具体区分。对于编码中带有二进制字符串或二进制矩阵形式的问题，可以将σ_{init}和σ_{final}分别对应于允许进行变异操作的变量的个数，也就是说让算法的执行过程从多点变异方式转换到单点变异方式。对于整数规划问题，可以仍然按照处理实数变量的方式执行算法，只是在产生新的变量值时采用向上或向下取整的方式进行。通过这些改进方式，学者们可以将IWO算法用于更多研究领域的问题求解过程中。当然，更多的算法扩展方式还有待于学者们去探索和应用。

结束语 本文对野草算法的研究及进展进行了较全面的论述。在短短的几年内，IWO算法引起了国际学术界的广泛关注，成为国际优化计算领域的研究热点之一，也逐渐成为解决实际工程优化问题的热门工具。

IWO算法虽然已对多种优化问题取得了很好的求解效果，并在诸多应用研究领域取得了长足的发展，但远未达到成熟的阶段，还有很多问题值得深入地分析与探讨。比如，算法是否存在统一的参数？如果不存在，如何针对不同问题合理设置参数组合？

与其他优化算法一样，关于IWO算法的理论研究尚显不足。理论研究工作者可以考虑算法收敛性、收敛速度、鲁棒性等方面的数学证明以及算法执行过程改进等方面。总体看来，IWO算法扩展了智能优化算法的领域，提供了求解优化问题的新思路，具有巨大的科研价值和应用潜力。

参 考 文 献

[1] Mehrabian A R, Lucas C. A novel numerical optimization algorithm inspired from weed colonization [J]. Ecological Informatics, 2006, 1(4): 355-366

[2] Mehrabian A R, Yousefi-Koma A. A novel technique for optimal placement of piezoelectric actuators on smart structures [J]. Journal of the Franklin Institute, 2011, 348(1): 12-23

[3] Zhang X, Wang Y, Cui G, et al. Application of a novel IWO to the design of encoding sequences for DNA computing [J]. Computers and Mathematics with Applications, 2009, 57(11/12): 2001-2008

[4] Mehrabian A R, Yousefi-Koma A. Optimal positioning of piezoelectric actuators on a smart fin using bio-inspired algorithms [J]. Aerospace Science and Technology, 2007, 11(2/3): 174-182

[5] Mallahzadeh A R, Oraizi H, Davoodi-Rad Z. Application of the invasive weed optimization technique for antenna configurations [J]. Progress in Electromagnetic Research, 2008, PIER 79: 137-150

[6] Rad H S, Lucas C. A recommender system based on invasive weed optimization algorithm [C]// IEEE Congress on Evolutionary Computation. Singapore, 2007: 4297-4304

[7] Kostrzewa D, Josinski H. The Comparison of an adapted evolutionary algorithm with the invasive weed optimization algorithm based on the problem of predetermining the progress of distributed data merging process [M]. Advances in Soft Computing, Berlin, Springer, 2009, 59: 505-514

[8] Pal S, Basak A, Das S, et al. Linear antenna array synthesis with invasive weed optimization algorithm [C]// International Conference of Soft Computing and Pattern Recognition(SOCPAR). 2009: 161-166

[9] Sahraei-Ardakani M, Roshanaei M, Rahimi-Kian A, et al. Study of electricity market dynamics using invasive weed colonization optimization [C]// IEEE Symposium on Computational Intelligence and Games. Perth, 2008: 276-282

[10] Mallahzadeh A R, Es'haghi S, Hassani H R. Compact U-array MIMO antenna designs using IWO algorithm [J]. International Journal of RF and Microwave Computer-Aided Engineering, 2009, 19(5): 568-576

[11] Mallahzadeh A R, Es'haghi S S, Alipour A. Design of an E-shaped MIMO antenna using IWO algorithm for wireless application at 5.8 GHz [J]. Progress In Electromagnetics Research, 2009, PIER 90: 187-203

[12] 苏守宝，方杰，汪继文，等. 基于入侵性杂草克隆的图像聚类方法[J]. 华南理工大学学报：自然科学版，2008，36(5)：95-105

[13] 苏守宝，汪继文，张玲，等. 一种约束工程设计问题的入侵性杂草优化算法[J]. 中国科学技术大学学报，2009，39(8)：885-893

[14] Zhang X, Wang Y, Cui G, et al. SIWO: A hybrid algorithm combined with the conventional SCE and novel IWO [J]. Journal of Computational and Theoretical Nanoscience, 2007, 4(7/8): 1316-1323

(上接第4页)

[13] Kumar R, Yao D D, Bagchi A, et al. Fluid Modeling of Pollution Proliferation in P2P Networks [C]// Proceedings of SIGMETRICS/Performance'06. 2006: 335-346

[14] Shi C, Han D, Hu X, et al. A unified model of pollution in P2P networks [C]// Proc. of Parallel and Distributed Processing(IPDPS). Apri 2008

[15] Kevin W, Sirer E G. Fighting Peer-to-Peer SPAM and Decoys with Object Reputation [A]// Proc. of the 2005 ACM SIGCOMM workshop on Economics of P2P systems [C]. Philadephia, Pennsylvania, USA, 2005

[16] Cristiano C, Vanessa S, Jussara A, et al. Fighting Pollution Dissemination in Peer-to-Peer Networks [A]// Proc. of the 2007 ACM symposium on Applied Computing [C]. Seoul, Korea, 2007

[17] Cristiano C, Jussara A. Reputation Systems for Fighting Pollution in Peer-to-Peer File Sharing Systems [A]// Proc. of the 7th IEEE Internation Conference on P2P Computing (P2P 2007) [C]. Galway, Ireland, 2007, 10

[18] Fabrici , B, Cristiano C, Marisa V, et al. Impact of Peer Incentives on the Dissemination of Polluted Content [A]// Proc. of the 2006 ACM symposium on Applied Computing [C]. Dijon, France, 2006

[19] Hong R J. Fault-Tolerant Mechanism for Removing Polluted Files in Peer-to-Peer Networks [D]. National Cheng Kung University, 2007

[20] Wei D, Yang S B, Guo L T. Design and Simulation of P2P File Sharing Anti-pollution System [J]. Journal of System Simulation, 2007, 19(24): 5705-5709