doi:10.11959/j.issn.1000-436x.2016207

内容共享网络中的关键问题

鲁强，刘波，胡华平
（国防科学技术大学计算机学院，湖南 长沙 410073）

摘 要：作为文件、信息和资源等联网内容的共享平台，内容共享网络在互联网中具有十分广泛的分布。随着网络技术的快速发展和深入应用，特别是对等网络的兴起与流行，极大地方便了人们通过网络分享各种内容。然而，大量的恶意文件、欺诈信息和间谍软件等不良内容蜂拥而入，对内容共享网络构成了日益严重的安全威胁。结合内容共享网络的发展与研究现状，从网络安全的视角对其概念内涵、类型划分、重要技术和发展与研究趋势等关键问题进行了综述。
关键词：内容共享网络；网络安全；测量技术；监测技术；抑制技术；发展趋势
中图分类号：TP393 **文献标识码**：A

Some critical issues of content sharing network

LU Qiang, LIU Bo, HU Hua-ping
(College of Computer, National University of Defense Technology, Changsha 410073, China)

Abstract: As a sharing platform for files, information and resources, CSN (content sharing network) possessed a very wide distribution in Internet. It did provide a great convenience for people to share various contents through Internet with the rapid development and wide application of network techniques, especially the emergence and prevalence of P2P (peer-to-peer). However, numerous malicious files, cheating information and spywares swarmed into Internet, which pose a serious threat to the security of CSN. A survey was made of some critical issues of CSN combining with the current development and research status from the perspective of network security. The critical issues include the definition and important properties, different taxonomies of types, key techniques, developing and researching trends of CSN.
Key words: content sharing network, network security, measurement techniques, monitoring techniques, mitigation techniques, developing trend

1 引言

计算机网络与信息技术的快速发展，不仅对国家政治、经济和文化等领域产生了重要的影响，而且不断覆盖、渗透到人们日常学习、生活和工作的方方面面，从传统的基础设施、物联网等物理空间和数据传递、信息共享等内容空间，逐步扩大到影响人们思想观念与行为决策的认知空间和社会空间。然而，目前的网络安全形势堪忧，在人们共享各种文件、信息和资源等内容的网络中，暴力、色情、虚假广告等不良内容及其夹带的病毒、木马、僵尸程序等恶意代码趋于泛滥，对网络安全和网络服务的健康运转构成了严重的威胁。因此，网络安全日益成为人们关注的焦点[注1]。

内容共享网络源于 Internet，出现在早期人们对各种文件数据的共享需求，即最早的文件共享系统，随着网络应用和服务的不断增长，人们需要分享的内容不仅局限于文件，还包括各种信息、应用

注1： 首届国家网络安全宣传周（中央网络安全和信息化领导小组办公室等主办）. http://www.xinhuanet.com/politics/2014gjwlaqxcz/ index1.htm。

收稿日期：2016-03-02；**修回日期**：2016-08-30
基金项目：国家自然科学基金资助项目（No.61572513）；高校博士点科研基金资助项目（No.20134307110016）
Foundation Items: The National Natural Science Foundation of China (No.61572513), The Education Ministry Doctoral Research Foundation of China (No.20134307110016)

和权限等新型内容。2002 年就有研究指出，明显带有暴力、色情色彩的 32 个关键词所对应的共享内容，在 eMule [注2]网络中就达到其内容总量的 2.5%，并且对此类内容发起搜索请求的用户占其用户总数的 5%，而这还只是当时可以统计到的冰山一角[1]。中国国家互联网应急中心近几年（2010 年~ 2014 年）的年度报告指出，我国境内感染木马和僵尸程序的主机由 2011 年的不足 900 万台到 2012 年一跃突破 1 000 万台，手机等移动智能终端设备感染量也逐年大幅上升，内容共享安全之形势由此可见一斑[注3]。

在人们日益增长的网络内容共享需求之下，传统 C/S (client/server)服务模式瓶颈凸显，P2P 技术应运而生，在互联网中获得广泛应用[2~6]。据国外某网站统计，Internet 上超过 50%的数据下载和 80%以上的数据上传都是借助于 P2P 网络来完成的[注4]。由此可见，P2P 内容共享网络代表着内容共享网络的发展趋势与潮流。另外，P2P 网络，尤其是结构化 P2P 网络的无中心特性，导致难以对其中传播的不良内容进行有效的监测和抑制，因此，本文将其作为重点研究对象。

目前，共享内容安全已经引起网络安全领域一些研究人员的关注[7~9]。相关的国际知名学术会议 USENIX-security、S&P (IEEE symposium on security and privacy)、CCS (ACM conference on computer and communications security)、LEET [注5](USENIX workshop on larger-scale exploits and emergent threat)、CRYPTO (advances in cryptology)、ICICSec (international conference on information and communications security)等也都将共享内容安全作为研究重点之一。

然而，内容共享网络的发展出现了不少新情况，如节点的自部署、移动化和社交化等，其概念内涵、类型划分、重要技术、发展与研究趋势等关键问题也不断扩展与丰富。

2　概念内涵

2.1　内容共享网络定义

关于内容共享网络的概念，目前没有统一、规范的认识和定义。由于 Internet 出现初期文件共享的率先兴起和其实例的广泛存在，人们通常会将内容共享网络默认为文件共享网络[10]。少量文献中也将其称为信息共享网络[11]、资源共享网络[12]。随着各种新型联网内容和特殊网络形式的出现，内容共享网络的概念内涵也不断延伸，例如共享载体不再局限于计算机主机，共享的网络也不再局限于传统的 Internet 等。内容共享网络概念延伸的同时，面临的安全威胁也愈演愈烈，如自 2011 年 CSDN 社区信息泄露以来，一些大规模网站相继出现的“拖库”“撞库”“洗库”等现象[注6]，以快播为代表的音视频共享网络引发的涉黄、侵犯版权等案件，社交共享网络中热传的雾霾视频、童星落榜等迷惑性强的内容，2014 年，仅新出现的恶意软件就多达 3.17 亿种，目前，恶意软件总量更是超过 20 亿种[注7]。为加强对安全的理解，下面给出内容共享网络的定义。

定义 1　内容共享网络。源于 Internet 出现早期人们对各种文件数据的共享需求，借助于网络并随着网络模式、技术等不断发展丰富，用于用户节点之间共享各种文件、信息和资源等联网内容而形成的一种社会关系网络。

内容共享网络 CSN(content sharing network)根据定义，包含以下 4 个要素。

1) 用户节点(UN, user node)。内容共享用户可以看作是形成网络结构的一个个节点。

2) 节点关联(NC, node correlation)。用户节点之间通过共享文件、信息和资源等内容形成一定的关联关系。

3) 内容的共享(CS, content sharing)。同一用户节点对于不同的内容，不同用户节点之间对于相同的内容，其共享程度都是有区别的，如公开或部分公开等。

4) 节点之间的交互活动(IA, interaction activity)。不同用户节点之间，除了内容共享之外，还会有其他社会交互活动，以维持内容共享网络的动态平衡性。

因此，内容共享网络可以记为 $CSN=(UN, NC, CS, IA)$。其中，$UN=\{UN_1, UN_2,\cdots,UN_K\}$，代表形

注2：eMule project, http://www.sourceforg.net/, 2005。

注3：CNCERT 2010/2011/2012/2013/2014 年互联网网络安全态势综述. http://www.cert.org.cn/publish/main/46/index.html。

注4：http://www.freemusictodownload.eu/p2p-statistics.html。

注5：LEET 是 2008 年由 WORM (ACM Workshop on Recurring/Rapid Malcode) 和 HotBots (USENIX Workshop on Hot Topics in Understanding Botnets) 合并而来。

注6：“拖库”指网站的用户信息数据库被入侵和窃取，“撞库”指利用从某网站获取的用户信息在其他网站进行测试以扩大用户信息的使用范围，“洗库”指将获取的网站用户信息变成现金等实际经济利益。

注7：Symantec website security solutions, http://www.symantec-secured.com, Symantec-WSTR-Whitepaper-APAC_PT1-SCN.pdf。

成内容共享网络的用户节点集，包括网络中本身存在的计算机、手机、虚拟机等普通节点，以及自部署的受控节点（含通过控制程序获取的恶意受控节点）；$NC=\{NC_{ij}\}(i\neq j$，两者都从 1, 2, …, K 中取值)，代表内容共享节点之间通过交互活动在网络结构中形成的关联集；*CS* 代表不同网络节点之间共享的内容集合；*IA* 代表不同网络节点之间的交互活动集合。

网络中每个用户节点可以用四元组 UN_i(*NodeID*, *NodeType*, *NodeRole*, *CSlist*)表示，即节点的标识、节点的类型、节点的角色、节点所共享的内容列表。其中，*NodeType* 指网络中共享节点的类型，包括正常节点、Sybil 节点（或者普通节点、自部署节点），*NodeRole*∈{*Normal*, *Guarding, Observing, Intercepting*, …}，是指节点在内容共享网络中所承担的角色，包括正常节点的一般角色和 Sybil 节点的守卫、监测和截流等角色，而 *CSlist* 为节点所共享的内容信息的链表结构。

节点之间的关联则可表示为四元组 NC_{ij}=(UN_i, UN_j, *CorDegree*, *IAlist*)，即关联节点 *i*、关联节点 *j*、节点之间的关联度、关联节点之间的交互活动列表，其中，*CorDegree* 的值初始默认为 0（节点之间没有交互活动且不是邻居节点）或 1（节点之间没有交互活动但互为邻居节点），并随着节点间交互活动的增多而变大，而 *IAlist* 为节点间所有交互活动信息的链表结构。

内容的共享用三元组链表结构 *CS*=(*ContentID*, *ContentType*, *ContentInfo*)表示，即共享内容标识、共享内容类型、共享内容信息，其中，*ContentInfo* 也是一个链表结构，包括共享内容的共享范围、共享程度等信息，还可以根据共享内容发展过程中出现的新变化进行相应的增加与修改。

节点之间的交互活动则用三元组链表结构 *IA*=(*IANum, IAType, IAInfo*)表示，即交互序号、交互类型、交互信息，其中，*IAInfo* 也是一个链表结构，包括交互的时间、交互的反馈等信息，可以根据交互活动中表现出来的新情况进行适当的扩展。

较之以往关于内容共享网络的概念和认识，本文不仅结合其近年来的发展变化给出了相应的定义，涵盖了僵尸网络[注8]、在线社交媒体等新的网络形式，而且较好地理解和考虑了内容共享网络的未来发展趋势。如定义中特别强调了内容共享网络的模式发展和社交化，突出了网络的拓扑结构和节点之间的社会关系；对网络中节点的类型和角色都进行了详细的区分，体现了节点的自部署和角色差异等新特点；共享内容和交互活动的强可扩展性，将会使内容共享网络的内涵更加丰富和全面。

2.2 内容共享网络关键性能

自内容共享的需求产生以来，其安全问题便如影随形，两者密不可分。从网络安全的角度而言，内容共享网络的以下性能备受关注。

1) 共享性。内容共享网络的最大特点就是共享，从一开始的满足文件的共享，发展到信息、资源的共享，在一定程度上适应了网络协同和节点协作的发展趋势，然而这给恶意代码等不良内容的“共享”也间接提供了相当有利的条件。随着共享内容的不断丰富，特别是主机、移动终端等硬件资源，一旦被恶意控制和利用，将会释放出强大的控制流量和破坏潜能。因此，对于网络的共享性这柄“双刃剑”，如何在减小乃至消除安全威胁的前提下，最大程度地发挥其在诸如共享、协同等方面的效能，是亟待解决的一大难题。

2) 透明性。透明性是指内容共享网络中的很多共享内容，共享与否以及共享范围和程度往往具有一定的默认设置，而对于用户节点来说则是透明的。另一方面，通常只有很少的用户节点会特别留意自己在网络上共享了哪些内容，以及这些内容的共享设置如何。内容共享网络应用的透明性，一般不提醒用户节点内容是否共享，或者只是提示用户节点的共享行为而不提供共享设置的更改等选项。透明性在给用户节点提供方便的同时，也逐渐暴露出内容共享网络在安全方面存在不少的隐患。

3) 传播性。传播性是指网络中的节点能够通过各种共享方式将特定的内容传递到网络中一定范围内的节点甚至整个内容共享网络。随着内容共享网络的社会化，各种共享内容特别是一些恶意内容的迅速传播且难以遏制，已经严重影响到当前网络共享环境的健康与和谐。除了传统的恶意植入与欺骗渗透，共享内容的传播还出现了一些新的模式，如社交网络媒体中用到的自媒体、朋友圈等[13]。内容共享网络的传播特性本身是为了更好地满足节点之间的共享需求，然而恶意内容的肆意泛滥让网络安全研究人员必须给予足够的重视。

4) 隐韧性。隐韧性一方面是指共享节点之间的组网、通信等活动很难被非授权节点发现，即隐蔽

注8：Know your Enemy: Tracking Botnets.http://www.honeynet.org/papers/bots.pdf。

性；另一方面是指内容共享网络在面临部分共享节点失效和恶意节点、内容持续渗透的情况下，仍然能保持一定共享效率的特性，即坚韧性。隐蔽性与坚韧性是紧密相联的，隐蔽性的提高可以降低共享节点失效的概率，反之坚韧性的提高则可以弥补隐蔽性要求对共享效率所产生的影响，二者相互促进，共同提高内容共享网络的生存能力。

3　类型划分

本节从网络模式、共享内容和节点组成等不同维度划分内容共享网络的类型，如图 1 所示。从网络模式的角度，将内容共享网络划分为非 P2P 内容共享网络和 P2P 内容共享网络 2 种类型；从共享内容的角度，根据网络主要共享的是哪种内容，将其划分为文件共享网络、信息共享网络和资源共享网络；从节点组成的角度，根据内容共享网络是否具有中心服务器节点，将其划分为集中式、非集中式和混合式。

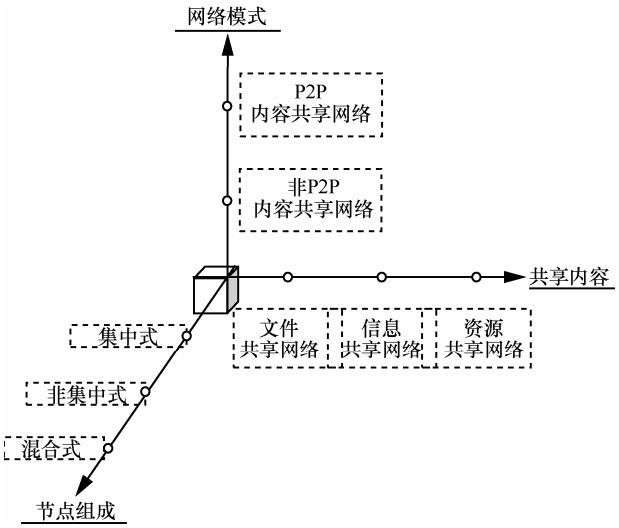

图 1　内容共享网络的多维度类型划分

3.1　基于网络模式的划分

由于内容共享网络产生的历史原因，一般的文献都会自然地将其根据网络模式的不同划分为非 P2P 内容共享网络和 P2P 内容共享网络[10,14,15]。另外，目前，大多数文献都比较关注 P2P 内容共享网络，因为非 P2P 内容共享网络本身的限制比较多（如节点 Web 访问数量上限、节点在线时间约束、共享内容过于依赖服务器等），相关的应用也比较少，而且文件共享等相关应用是 P2P 网络的优势，P2P 网络模式又代表着内容共享网络的发展趋势与潮流。

非 P2P 内容共享网络的典型代表是 C/S 共享网络，包括传统的大型企业、单位 Web 共享网站，客户端程序，以及个人通过蓝牙、Wi-Fi 等新兴网络技术共享计算机、移动设备上的文件等内容。

P2P 内容共享网络，顾名思义就是采用了 P2P 协议的内容共享网络，克服了非 P2P 内容共享网络的一些限制，如节点之间的直接共享打破了 Web 访问量上限，节点的在线时间也更加自由，共享内容的分散存储降低了大规模内容共享的服务器依赖等。P2P 模式本身是 Internet 出现之初分布、开放共享的预期设计，后期由于网络节点的不对称性发展和共享所需网络等硬件资源的限制等原因，没有得到很好的发展。应用促进发展，自 1999 年 P2P 的第一个经典共享实例 Napster[16]的出现，P2P 文件共享等应用广泛兴起，极大地带动了 P2P 内容共享网络的发展，使其网络模式的主流重新回到了 P2P 模式。

3.2　基于共享内容类型的划分

根据共享内容的类型，内容共享网络划分为文件共享网络、信息共享网络、资源共享网络。这种类型的划分，主要针对的是共享网络各种实例中所共享的内容，哪种共享内容占主导地位，那么共享网络就属于相对应的网络类型。

文件共享网络是最早出现的，也是目前应用最为广泛的。Internet 出现早期，人们的共享需求主要就在于文件，并且文件共享这一需求在网络节点的内容共享活动中一直占据着重要的位置，在 P2P 内容共享网络中更是得到了广泛的应用。文件共享网络比较典型的实例是共享音乐文件的 Napster、共享种子文件的 BitTorrent[17]、共享多媒体文件的 Gnutella[18]等。文件共享网络的发展趋势是不局限于某种特殊的文件，而是共享各种类型的混合式文件，从传统的数据文档，到网络流行的音、视频流，以及图片、应用程序等，如 eMule[19]不仅共享了所有这些类型的正常文件，甚至还出现了这些之外的恶意代码和不健康视频等。

信息共享网络，最典型的就是僵尸网络，各僵尸节点之间共享各种通信、命令等信息。僵尸网络，是攻击者通过一对多的控制结构，恶意组织大量受控网络节点形成的受控网络[20~22]。因此，为了触发相应的注入、传播或攻击等活动，各僵尸节点之间就要共享相应的指令信息。另外，

除了僵尸网络，一些大型门户网站与用户节点，以及计算机、移动设备的客户端节点等，形成的也是信息共享网络。

资源共享网络，主要是指通过网络共享主机、移动终端、传感器等各种设备的时间和空间资源，如 CPU 处理周期、进程、存储空间、带宽等。目前非常典型的资源共享网络不是很多，利用 LOIC [注9]等开源平台中的主机资源形成的自部署内容共享网络可以算作一个。

通常情况下，内容共享网络在共享内容方面的类型划分并没有非常严格的界限，文件、信息和资源之间存在着一定的交叉，或者说文件共享网络、信息共享网络和资源共享网络三者可以看作内容共享网络发展过程中出现的名称不同、本质相同的 3 个概念。

3.3 基于节点组成方式的划分

目前，基于节点组成的类型划分，主要都是针对 P2P 内容共享网络，如文献[15]将内容共享网络分为中心式、无结构、结构化和混合式，类似于对一般 P2P 网络结构的划分，而没有考虑非 P2P 内容共享网络；文献[23]针对僵尸网络等特殊的内容共享网络，按节点组成划分为单服务器网络、多服务器网络、分层网络和不规则网络，其中，不规则网络则主要是指 P2P 内容共享网络，但是大部分多服务器网络和分层网络通常也是 P2P 内容共享网络，如 KaZaA[24]、QVOD [注10]等；文献[25]将内容共享网络分为中心式、非中心式和随机式，但是有很多随机式共享网络实质上是中心式的。综合以上的分类，本着更准确的分类原则，本文根据节点组成是否具有中心服务器，将当前的内容共享网络分为集中式、非集中式和混合式。

集中式内容共享网络，存在明显的中心服务器，可能是单服务器，也可能是多服务器[26]。集中式内容共享网络包括 Web 内容共享网络和部分 P2P 内容共享网络，其采用的通信协议包括 P2P、HTTP[27]、FTP[28]、IRC[29]，以及各种即时通信[30]、邮件协议[31]等，典型实例有 Napster、BoBax[32]、BitTorrent、FS2You[33]、eDonkey2000[34]、iKee.B[35]等。集中式内容共享网络的明显优势在于控制性好，主要不足是面临单点故障的威胁[36]。

非集中式内容共享网络，摒弃了中心服务器，克服了单点失效的威胁。此类内容共享网络采用的通信协议为 P2P，包括结构化的 P2P 和非结构化的 P2P。结构化 P2P 内容共享网络的典型实例有 Chord[37]、CAN[38]、Tapestry[39]、Kademlia[40]和 Storm[41]等，而非结构化 P2P 内容共享网络有 Gnutella、Freenet[42]等。非集中式内容共享网络的主要优势是隐韧性高，不足之处是遭受 Sybil 攻击[43]的威胁更大。

混合式内容共享网络是指在节点组成上兼具集中式和非集中式 2 类网络特点的内容共享网络。例如，KaZaA 中引入了超级节点，普通节点和所属的超级节点之间形成了一层集中式组网结构，所有超级节点之间又是非结构化的 P2P 组网结构[24]；Waledac 被发现是一种复杂的 4 层混合式信息共享网络[44]；eMule 中的普通节点不仅与所属 eMule 服务器形成了一层覆盖网络（overlay network [注11]），而且还与 eDonkey2000 服务器也形成了另外一层覆盖网络，并且这 2 层覆盖网络之间通过普通节点的连接使整个 eMule 网络更具扩展性，隐韧性也更好[45]；QVOD、迅雷看看等新兴音、视频在线点播网络平台的用户节点组成本身是非集中式的，而与网站、索引、数据等多类服务器之间则形成了集中式的共享关系。混合式内容共享网络旨在综合集中式和非集中式这 2 种节点组成的优点，但是面临的威胁也不可避免地更多。

3.4 典型内容共享网络的类型分析

表 1 所示是一些典型内容共享网络在网络模式、共享内容和节点组成等维度上的分类情况。从表 1 中也可以看出，P2P 文件共享是当前内容共享网络中应用最为广泛，混合式的节点组成是内容共享网络结构发展的趋势之一。

4 重要技术

当前对内容共享网络的研究主要集中在 3 个方面：1) 特点规律性研究，主要研究内容共享网络的概念、应用发展、网络模式、节点组成等方面的特点和规律；2) 安全防护类研究，主要研究如何应对内容共享网络中出现的各种不良内容，包括不良内容的监测与抑制技术等；3) 控制利用型研究，主要研究如何高效地利用内容共享网络的共享、传播和

注9：Low orbit ion cannon. https://en.wikipedia.org/wiki/Low_Orbit_Ion_Cannon。

注10：QVOD(快播), quick video on demand. http://www.qvod.com。

注11：Overlay network. https://en.wikipedia.org/wiki/Overlay_network。

表 1　　**各种内容共享网络的分类情况**

典型内容共享网络	按网络模式分类		按共享内容分类			按节点组成分类		
	非 P2P 内容共享网络	P2P 内容共享网络	文件共享网络	信息共享网络	资源共享网络	集中式	非集中式	混合式
C/S	√		√			√		
Napster		√	√			√		
eDonkey2000		√	√			√		
Gnutella		√	√				√	
BitTorrent		√	√			√		
eMule		√	√					√
LOIC		√			√		√	
iKee.B	√			√		√		
FS2You		√	√			√		
KaZaA		√	√					√
Waledac		√		√				√
Storm		√		√			√	
QVOD		√	√					√

社交等特性，实施特定内容的大规模传播和对目标节点集的多空间、连锁式影响等。

由于本文以增强内容共享网络的安全为研究目标，因此本文的重要技术主要集中在前 2 个方面。如图 2 所示，重要技术具体又可以分为网络测量技术、不良内容监测技术和不良内容抑制技术 3 类。其中，网络测量技术包括主动测量和被动测量；不良内容的监测技术包括针对网络特性的监测和基于 Sybil 节点的监测；不良内容的抑制技术包括对其索引的污染和传播路径的破坏。

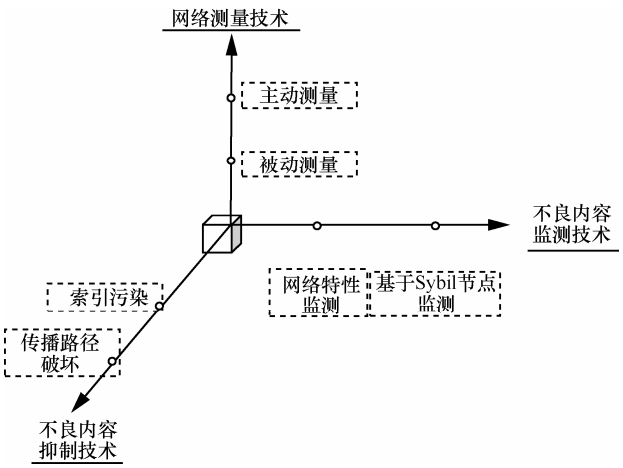

图 2　重要技术分类

内容共享网络测量与不良内容的监测和抑制的 3 类技术之间联系紧密，其关系如图 3 所示。网络测量为不良内容的监测与抑制提供网络拓扑、节点渗透依据等必要的基础与支撑，不良内容的监测与抑制则为网络测量提供相应的补充；不良内容的监测为其抑制提供索引与传播路径等关键信息，从而实现监测的目的。

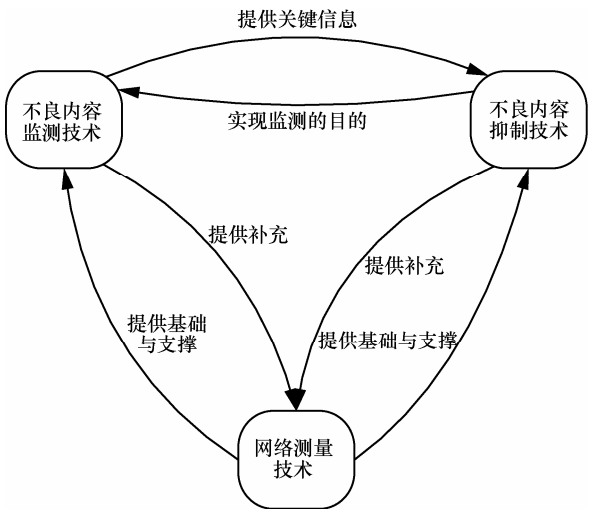

图 3　重要技术之间的关系

4.1　网络测量技术

目前，对于各种网络的测量，都可以分为主动测量和被动测量，内容共享网络也不例外，其测量技术亦可从主动和被动这 2 个方面来分析。

4.1.1　主动测量技术

主动测量是指主动向内容共享网络发送探测

报文如请求报文，并通过记录和分析网络节点的响应报文来获得网络的节点情况及其连接关系。由于早期的 P2P 内容共享网络多数建立在免费、开源的基础之上，使研究人员可以根据相应的协议规程进行某些细节上的修改，从而得到各种不同的定制版网络测量客户端。后来由于版权问题和 P2P 的逐步商业化，使主动测量受到很多限制，但是研究人员仍然可以通过逆向工程技术，对私有协议进行一定层次的解析。

主动测量方面的典型实例比较多，如美国国家科学基金早期资助的 AMP (active measurement project)研究项目[46]；P2P 出现早期 Stefan 等[47]通过对 Gnutella 和 Napster 的拓扑测量，发现 Gnutella 的节点在连通度上呈现出 Power-law 规律，后来 Ripeanu 等[48]通过测量又对这一规律进行了一定的修改和补充，而 Stutzbach 等[49]则通过改进的主动测量揭示了 Gnutella（研究平台为 LimeWire 和 BearShare）中不符合上述规律的超级节点，其连接度为某些特定的常数（如 30、45 等）；KAD 爬虫 Blizzard 在获取共享节点数量方面效率优势明显，可以在一个小时左右获取 150 万~300 万节点[50]；Steiner 等[51]通过对网络中的超级节点 ID 与普通节点 ID 的测量，发现了两者对内容共享网络整体的影响力有着显著的差异，但是测量在获取节点关系、网络拓扑等信息方面，有效性较差；余杰等[52]通过对结构化 P2P 内容共享网络的主动测量，分析发现了网络中存在节点 ID 重复、消息无验证等安全缺陷。

4.1.2 被动测量技术

被动测量与主动测量不同，它不会向网络发送数据报文，信息的获取源自网络节点设备的运行日志文件或者网络链路、设备固有的流量记录等，所以不会对网络运行本身造成额外的负担。被动测量技术一般适用于大规模的网络测量，其测量的重点在于 P2P 流量识别。

被动测量方面的实例相对较少，如美国国家科学基金早期资助的另一研究项目 PMA (passive measurement and analysis)[53]；Huang 等[54]通过在骨干网络部署探测点，利用被动测量方法对网络的运行错误进行分析和定位。

4.1.3 技术比较与分析

主动测量的优点是数据采集准确，数据量小，便于后期处理，且该数据集能够直接揭示各用户节点的网络行为。缺点在于，测量过程中会给网络引入额外的流量负担，同时大范围的测量对测量发起点的数量和带宽又有很高要求，而且从网络边缘进行探测，用户的连通性也是很大的一个问题，用户集很难做到完备。

被动测量的优点则是数据记录相对完备，不会给网络增加负载，且通用性较好，可以用来分析多种 P2P 系统。但是，其缺点也是显而易见的，数据量大，需要对采集数据进行过滤筛选和处理，数据识别不精确，而且往往无法提供像主动测量那样直接且针对性强的测量结果。另外，被动测量需要在网关部署测量系统，难以实现对全球 P2P 网络活动的有效测量。

内容共享网络测量在节点枚举和结构探测上，可以为 Sybil 节点的高效渗透提供相对准确、全面的网络拓扑信息，同时也对网络测量的实时性提出了较高的要求。然而，单纯的被动测量很难满足这一要求；现有的主动测量方法虽然节点枚举速度快，但没有实现对网络拓扑结构的高效探测。因此，如何快速、准确地获取大规模内容共享网络的结构信息是一个亟待突破的技术难点。

4.2 不良内容监测技术

4.2.1 网络特性监测技术

通过对共享网络特定性能的监测来发现网络的异常，进而判断是否有不良内容的传播活动，是不良内容监测研究最基本的思路。一般来说，这种监测技术可以为发现不良内容提供某些参考和判别的规律，但是往往通用性、针对性差，无法满足大规模内容共享网络全覆盖监测的需求。

例如，Christin 等[55]以网络内容的可用性为依据，对 4 种流行的内容共享网络（Gnutella、eMule、Overnet 和 FastTrack）进行了相应的监测，并根据查询响应比例与时间、内容重复性和下载效率等指标进行了深入的分析。其中，“内容重复度在分布上基本呈幂律规律，且内容的重复度越高，节点的下载效率也越高”这一规律可作为判断网络中内容分享异常与否的参考。René Brunner 等则利用经过改进的 aMule [注12]客户端，监测 KAD 网络节点的可用性、内容的搜索与发布效率等特性，以分析、辨别节点的共享行为[56]。更进一步，Steiner 等[57,58]对 KAD 网络全局进行了 6 个多月的持续监测，发现了网络在节点的平均数量、地域分布、交互时间

注12：AMule. http://wiki.amule.org/wiki/AMule。

间隔等方面存在特定的规律，可借鉴于发现不良内容传播等异常事件。

4.2.2 基于 Sybil 节点的监测技术

基于 Sybil 节点的监测，是在网络特性监测技术的基础上利用 Sybil 节点感知网络共享活动的特性而形成的一种新的监测技术。

Holz 等[59]首先提出了一种基于 Sybil 节点的 P2P 网络监测思路。该思路主要通过将大量 Sybil 节点加入到 KAD 网络，以持续感知网络中其他节点的搜索活动情况来达到对全网实施监测的目的。Wang 等[60]则进一步通过建立预测模型和开展模拟实验，分析和讨论了不同比例的 Sybil 节点对基于 Kademlia 协议的共享网络监测效果的影响。Lu 等[61]在确保 Sybil 节点比例最优的情况下，研究了利用多种不同角色的 Sybil 节点共同对内容共享网络中的不良传播内容进行监测。

4.2.3 技术研究现状分析

总之，目前关于不良内容监测技术的研究较为缺乏：一方面是缺乏对大规模内容共享网络中共享活动突出特性的发掘与运用，如当前网络中出现的不良内容在热点内容中所占比例不断攀升等新特性；另一方面是缺乏有效提高 Sybil 节点监测能力的技术方案，在 Sybil 节点的研究上量要多于质，Sybil 节点的重要性划分及其之间的协同问题鲜有研究。

以典型的内容共享协议 KAD 为例，目前在协议改进之后，同一 IP 地址难以产生很多不同的 Sybil 节点来达到同时覆盖 KAD 网络中的大量节点；另外，由于 KAD 节点 K 桶分布特点，单个 Sybil 节点很难加入大部分 KAD 节点的路由表。这些都显著削弱了 Sybil 节点对 KAD 网络的监测能力。

4.3 不良内容抑制技术

4.3.1 技术概述与分析

目前，不良内容抑制技术相关的研究，大致可以概述为以下 5 个方面.

1) Sybil 攻击。Douceur 等[43]指出，在没有中心认证机制的 P2P 匿名共享系统中，要想完全防御 Sybil 攻击是不可能的。Carlton 等[62,63]则更进一步，对 Sybil 攻击破坏僵尸网络 Storm 节点共享信道的效果展开了定量研究：他们模拟在 Storm 网络中加入大量的 Sybil 节点，每个 Sybil 节点接收到来自 Storm 节点的任何搜索请求时，都返回错误的应答消息，使 Storm 节点对命令与控制信息的搜索失败；分析结果表明，这种 Sybil 节点可以在较大程度上达到抑制 Storm 共享网络信道的目的，而无需对命令与控制信息的发布 Key 值进行预测和分析。

2) 假块污染。假块污染是指发起者通过伪造大量虚假客户端加入到内容共享网络中，这些虚假客户端一旦接受网络中其他节点的下载请求后，就会提供虚假的数据上传。这样，下载节点会由于校验失败而丢弃下载到的数据分块，并重新下载。虚假客户端还会通过提高被其他节点请求的几率来达到占用其他节点的下载带宽、减低其下载速度的目的。从本质上来说，这也可以看作是一种 Sybil 攻击[64]。为了使大量节点都能够下载到假块，假块污染对网络和存储资源有较高的要求。

3) 索引污染。索引用于帮助网络用户定位目标内容的存储位置。索引污染的发起者通过向网络中发布大量虚假索引信息来阻止用户正确获取目标内容[65]。虚假索引信息往往指向错误的网络地址或端口[66]；而当用户试图与虚假索引指向的网络地址建立连接时，一般都会失败。索引污染与假块污染相比，发起者无需向请求者传送文件，所需带宽等服务资源少。因此，在各种方法中，索引污染更加低耗、高效[67]。

4) 拒绝服务。发起者通过持续不断地连接到目标内容所在节点，极大耗费该节点的上传带宽，从而阻止网络其他用户从此处下载文件[68]。

5) 路由表污染。发起者通过修改、劫持正常网络节点的路由表项，使它们无法与其他正常的网络节点进行通信。这样，发起者就能够假装是正常节点所要寻找的目标节点，任意构造信息返回给这些发起查询的节点，一定程度上控制查询节点的网络行为[69]。

假块污染和拒绝服务对发起者的网络和存储资源都要求较高，而当前很多内容共享网络从协议层面增强了安全性设置（如 eMule 网络的节点在路由表满时无法加入新的虚假邻居节点，且对路由表中的邻居节点定期进行检查以防止来自同一网络地址的多个邻居节点占据路由表项等），使路由表污染常常难以奏效。因此，相对而言，综合考虑成本和效果，Sybil 攻击技术和索引污染技术更适于对大规模内容共享网络中不良内容的传播活动

进行抑制。然而，目前，Sybil 攻击技术中各个节点之间协同抑制的潜力尚未被有效挖掘；索引污染技术也需要从污染策略上进行更多的研究，以提高抑制效能。

这些抑制技术，不论适用与否，效能如何，从方法上都可以进一步归纳为索引污染和路径破坏（Sybil 攻击、假块污染、拒绝服务、路由表污染）两大类，如表 2 所示。

表 2 不良内容抑制技术比较

方法类型	抑制技术	适用	效能
索引污染	索引污染	低耗高效	抑制策略
		相对适用	需要提升
路径破坏	Sybil 攻击	低耗高效	协同潜力
		相对适用	有待挖掘
	假块污染	对网络和存储资源要求较高	
	拒绝服务	对网络和存储资源要求较高	
	路由表污染	协议安全设置使其常常难以奏效	

4.3.2 索引污染技术

在内容共享网络中，将存储特定标识与相应内容的对应关系（即索引信息）的节点，称作根节点。对于其他的节点，要得到目标内容，首先需要对特定的标识进行搜索，从根节点获取相应的索引信息，进而获取目标内容。在分析得到内容共享网络中不良内容的标识之后，如若能够改变根节点上的这种对应关系，便可使其他节点难以获得不良内容。不良内容的索引污染即是研究如何改变根节点中与特定标识对应的索引条目，以此有效抑制不良内容的传播。

内容共享网络往往具有“搜索—复制”的特性，即搜索节点在获得不良内容的索引信息后，会进一步向其他正在搜索该信息的网络节点提供该索引，因此，索引污染研究必须要和不良内容的传播路径破坏研究结合起来，共同形成互补体系，方可实现高效抑制。

4.3.3 路径破坏技术

内容共享网络中，不良内容的传播路径包括 2 部分：1) 不良内容发布者将信息发布到内容共享网络的发布路径；2) 网络中其他节点获取不良内容的搜索路径。不良内容的传播路径破坏研究如何干扰发布者的发布活动与尝试获取其他节点对不良内容的搜索活动，使不良内容无法正常传播到网络节点，从而有效抑制不良内容的传播。

5 发展与研究趋势

本节在前面几节概述与分析内容共享网络基本概念、类型划分和重要技术等关键问题的基础上，对内容共享网络自身的发展与研究趋势进行了概括与提炼。其中，发展趋势主要是对内容共享网络自身特点和性能发展方向的预计，研究趋势则主要是对内容共享网络研究未来所面临的技术难点和焦点的推测。

5.1 内容共享网络发展趋势

内容共享网络在与各种不良内容的博弈过程中，为了实现健康共享与传播等目标，不断地发展演化，并表现出以下发展趋势。

1) P2P 和混合化。内容共享网络在网络模式上，呈现出 P2P 化趋势；在节点组成上，则朝着混合式发展。内容共享网络的 P2P 化日益明显，越来越多的内容共享网络在网络模式上倾向于选择 P2P 组网模式。内容共享网络的 P2P 化，代表着非集中式的节点组成。然而，集中式和非集中式内容共享网络自身都存在明显的缺陷，混合式内容共享网络将二者结合得以发展，并且随着 eMule 等典型实例的广泛应用而逐步发展成为节点组成的趋势。

2) 小型化。通过长期对 eMule 等典型内容共享网络的测量分析，本文发现网络的在线节点数量相比前几年有缩减之势[50,61]。如 eMule、BT 等网络的同时在线节点数量达到前几年那样百万量级的时候很少，但是网络的规模能够稳定在 10^5 这一量级。因此，这里所说的小型化不是指内容共享网络规模会越来越小，而是随着网络带宽、安全防护等因素的变化而逐步缩减至相对稳定的某一量级。硬件的持续更新和带宽的不断增加，使更小规模的内容共享网络就能满足网络节点的各种共享需求。另外，自部署内容共享网络的出现，也标志着小型化的内容共享网络在灵活性和效率上的优势正在逐步得到体现。

3) 社交化。作为一种社会关系网络，内容共享网络的社交化程度本身就要比一般的网络高，并且随着在线社交网络的兴起与发展，其社交化特性也越来越明显[70,71]。共享内容的日益丰富和 UGC（user generated content）[注13]发展模式，使网络共享的层次逐步从物理、信息空间扩展到认知与社会空间，其

注13：UGC. http://baike.baidu.com/subview/713949/9961909.htm。

影响范围也不再局限于单纯的网络与信息领域，而是不断向人们的心理和意识领域渗透。

4) 移动化。对于内容共享网络，移动化为其提供了新的发展平台，越来越多的移动设备在内容共享网络中得到应用[72~74]。从早期的诺基亚 Symbian、微软 Windows Mobile，到当前流行的谷歌 Android、苹果 iOS 等，不仅通过移动客户端形成了各种移动化的内容共享网络，还出现了因安全漏洞被利用而形成的恶意内容共享网络，如恶意代码 Geinimi [注14] 感染的安卓手机构成的信息共享网络，以及臭名昭著的手机间谍游戏（愤怒的小鸟）用户节点所形成的“情报搜集”网络等。中国互联网络信息中心的统计数据显示，截至 2015 年 6 月，中国网民规模达到 6.68 亿人，其中，手机网民约为 5.94 亿人[注15]，充分说明了内容共享网络的移动化趋势。

5) 智能化。内容共享网络的移动化，往往伴随着智能化的发展。随着物联网和智能穿戴设备的兴起与发展，共享节点的智能性也不断得到提高，能够根据自身的兴趣和需求对共享内容进行定制。网络的智能化为用户节点共享各种内容提供着越来越多的便利，与此同时，其安全化的发展也必须得到足够的重视，如能够及时监测发现并有效抑制各种恶意共享内容的传播，营造安全健康的网络共享环境。

6) 安全化。由于共享是重心，加之 P2P 模式的开放、匿名等特性，内容共享网络初期在安全机制上比较缺乏，因而出现了诸如恶意共享软件、下载链接和文件泛滥等网络安全问题。在内容共享网络的发展过程中，这些安全问题逐步得到重视，也出现了一些安全化举措，如 BitTorrent、Waledac、Peacomm [注16]等共享网络中相继加入了一系列越来越复杂的信誉激励、传输加密、Fast-Flux [注17]传播隔离等机制和技术。安全性与共享性将会更加同步，而不再是安全问题的解决滞后于共享性提升的状态。

注14：Android. Geinimi, https://www.symantec.com/security_response/writeup.jsp?docid=2011-010111-5403-99。

注15：CNNIC《第 36 次中国互联网络发展状况统计报告》. http://www.cnnic.net.cn/hlwfzyj/hlwxzbg/hlwtjbg/201507/P020150723549500667087.pdf。

注16：A multi-perspective analysis of the storm (peacomm) worm. http://www.cyber-ta.org/pubs/StormWorm/SRITechnical-Report-10-01-Storm-Analysis. pdf。

注17：Fast flux. https://en.wikipedia.org/wiki/Fast_flux。

5.2　内容共享网络研究趋势

结合内容共享网络的相关研究与发展趋势，本文从特点规律性、安全防护类和控制利用型分析了内容共享网络的研究趋势，并对其未来研究所面临的技术难点和焦点进行了推测。

5.2.1　特点规律性研究

特点规律性研究方面的技术主要包括内容传播模型技术、网络协议解析技术和内容共享网络效能评价技术等。

1) 内容传播模型技术。共享内容的传播模型能够反映网络中节点的共享行为，对于理解内容共享网络的概念与本质，指导安全防护类研究等具有重要的基础意义。内容传播模型目前的研究主要集中在不良内容的病毒式传播，如文献[75]基于流行病学理论对 P2P 内容共享网络中被动式蠕虫传播进行了建模和分析，以准确标识节点的共享行为和预测蠕虫传播的趋势。但是，共享内容不仅有社交圈、网络媒体、邮件附件、网址链接、广告植入等多种传播方式，而且对于其他不良内容以及正常内容的传播模型的研究也较少，这些都有待深入。

2) 网络协议解析技术。内容共享网络的协议解析技术将节点的共享活动和交互关系进行形式化的表达，便于分析和理解网络共享行为的一些共性和特殊规律，为网络测量和不良内容监测与抑制提供重要的参考信息。目前，对于共享网络协议的解析大都还停留在人工分析的层次，缺乏相应的协议分析引擎，不能做到快速、准确的自动化解析。不良内容蜂拥入网，使用的协议也不断变化，形式和版本众多，普适性好、自动化程度高的共享网络协议解析技术乃至平台必将成为研究的趋势和热点之一。

3) 内容共享网络效能评价技术。内容共享网络的效能评价主要涉及安全和效率，基本原则是在提高安全性的同时能够不影响网络共享的效率。Li 等[76]指出，对于所添加的安全功能，要在增强网络安全的同时，不能明显影响网络本身的性能，或者对性能的影响能够控制在可容忍的范围之内。Liu 等[77]则初步提出了一些评价指标，如在网络管理方面涉及网络系统结构的维护、节点通信、内容搜索等网络活动的开销、效率方面则包含搜索等正常网络活动的成功率和时效等因素。然而，更加系统、通用的内容共享网络效能评价指标还有待研究，体系化和平台化将是研究的必然趋势。

5.2.2 安全防护类研究

安全防护类研究方面的技术可以大致归纳为不良内容全局监测技术、Sybil 节点协同监测与抑制技术和不良内容溯源和清除技术等。

1) 不良内容全局监测技术。对于不良内容的全局监测，需要立足互联网全局，将网络中所有的共享活动进行全方位、多方法、异策略的监测，以快速发现各种不良内容。内容共享网络的发展趋势，如 P2P 和混合化、小型化等，使原来基于关键节点、流量统计等相关监测方法的效率和成功率都呈现出下降趋势。因此，内容共享网络的监测应当从全局出发，综合利用多种方法和策略，克服不同内容共享网络的异构性，在节点的共享活动和行为上进行聚焦，从而形成对不良内容长期、有效的监测能力。

2) Sybil 节点协同监测与抑制技术。不良内容的监测和抑制中都有利用到 Sybil 节点的思路，但是由于安全设置和技术实现上的一些难点，Sybil 节点对于单独的监测和抑制还有待加强，基于 Sybil 节点的协同监测与抑制技术应该作为一个重要的研究方向。不良内容的监测与抑制，从内容共享网络安全防护类研究的角度来看，本身就是一个整体，完全孤立的监测或抑制所能够达到的安全效能肯定不是最佳的，两者之间的协同机制是研究的必然趋势。当然，对于各种不同角色的 Sybil 节点，如何形成有效的控制体系、具体的监测与抑制等任务的分工协作，以及 Sybil 节点监测与抑制各自难点的攻克等，都是非常重要的研究点。

3) 不良内容溯源和清除技术。在安全防护方面，监测和抑制是针对不良内容最主要的措施，但是一定的溯源和清除技术也是必需的。对于一些不良内容，如果在监测和抑制的基础上，加上溯源和清除等补充手段，不仅能够起到更好的效果，而且有助于对共享和传播本质及特性的深入理解。如何在监测和抑制的基础上做到溯源和清除，是安全防护类研究的重要方向之一。

5.2.3 控制利用型研究

控制利用型研究虽然不是本文的研究重点，但是也是内容共享网络研究的一个重要方向。对于内容共享网络的控制利用，虽然被恶意利用可能产生严重的危害，但是经过良好的引导可以为内容共享网络的特点规律和安全防护研究提供强有力的技术支撑。结合对内容共享网络的理解和分析，本文认为控制利用型技术主要包含以下 2 点。

1) 内容共享网络协同技术。内容共享网络的控制利用，在协同上区别于共享节点层面的分工协作，主要是指不同共享网络之间的跨网层次的协作。这种跨网协同的技术，在一定程度上符合了内容共享网络的小型化发展趋势，通过将多个异构的小网络组织、管理并且调度起来，实施大规模内容共享网络的功能和作用，有利于更多乃至全网共享资源的合理分配和充分利用。网络层面的协同技术相比节点层面的更为复杂，对共享性与透明性的要求更高，相应的研究非常必要和重要。

2) 内容共享网络高效控制技术。对内容共享网络的高效控制，是一个综合、复杂的技术集合，旨在提高网络的隐韧性和最大化特定共享内容的影响范围。高效控制不仅面向多类型、多角色和多变化的共享对象，需要处理共享活动的多层空间，而且必须能够形成包含最优解或多个次优解的通用、可验证控制策略集。因此，其挑战与机遇并存，是一个庞大并且重要的研究方向。

除了以上 3 个方面的技术难点和焦点之外，内容共享网络中相关的法律法规和道德准则在应对个人隐私和知识版权等问题上也发挥着不容忽视的作用，它们共同形成构建和谐、健康、绿色共享环境的重要因素。作为一种可控制、可利用的网络，内容共享网络理应有其不分国界的人类伦理边界和配套的国际法律依据，以避免内容共享网络产生严重的安全威胁[78]。

6 结束语

内容共享网络作为文件、信息和资源等联网内容的共享平台，其环境的安全和谐与否关系到国家和社会生活的多个层面，直接与潜在的安全影响都不容小觑。随着网络安全逐步上升为国家战略，内容共享网络中蜂群式的恶意文件、不良信息和间谍软件等安全问题已经成为网络安全领域的关注焦点和研究热点。本文在概述和分析内容共享网络的基础上，对内容共享网络进行了新的定义，剖析了其关键性能；给出了内容共享网络在网络模式、共享内容和节点组成等多个维度的类型划分，并分析、比较了一些典型的内容共享网络实例；详细介绍了内容共享网络在网络测量、不良内容监测技术和不良内容抑制等方面重要技术的研究现状与最新进展；探讨了内容共享网络自身的发展与研究趋势。

参考文献：

[1] STANIFORD S, PAXSOM V, WEAVER N. How to own the Internet in your spare time[C]//The 11th VSENZX Security Symposium. San Francisco, 2002: 149-167.

[2] CLARK D. Face-to-face with peer-to-peer networking[J]. IEEE Computer, 2001, 34(1): 18-21.

[3] MILOJICIC D S, KALOGERAKI V, LUKOSE R, et al. Peer-to-peer computing HPL-2002-57[R]. Palo Alto, USA: HP Laboratories, 2002.

[4] RATNASAMY S, KARP B, YIN L, et al. GHT: a geographic Hash table for data-centric storage[C]//The First ACM International Workshop on Wireless Sensor Networks and Applications. New York, ACM, 2002: 94-103.

[5] SCHODER D, FISCHBACH K. Peer-to-peer prospects[J]. Communications of the ACM, 2003, 46(2): 27-29.

[6] CHEN H, YANG M, HAN J Q, et al. Maze: a social peer-to-peer network[C]//The IEEE International Conference on E-Commerce Technology for Dynamic E-Business. Beijing, 2004: 290-293.

[7] JIA D M, YEE W G, FRIEDER O. Spam characterization and detection in peer-to-peer file-sharing systems[C]//ACM Conf on Inf and Knowl Mgt (CIKM). ACM, 2008: 329-338.

[8] WANG Q Y, VU L, NAHRSTEDT K, et al. MIS: malicious nodes identification scheme in network-coding-based peer-to-peer streaming[C]//The 29th Conference on Information Communications (INFOCOM 2010). Piscataway, NJ, USA, 2010: 296-300.

[9] SHI J T, ZHANG H L. A protocol based countermeasure to BitTorrent fake-block attack[J]. Journal of Computational Information Systems, 2012, 8(12): 5211-5218.

[10] SAROIU S, GUMMADI K P, STEVEN D G. A measurement study of peer-to-peer file sharing systems[C]//International Society for Optics and Photonics. 2001: 156-170.

[11] PARAMESWARAN M, SUSARLA A, ANDEW B, et al. P2P networking: an information-sharing alternative[J]. Computer, 2001 (7): 31-38.

[12] IAMNITCHI A I. Resource discovery in large resource-sharing environments[D]. Chicago, America: University of Chicago, 2003.

[13] TREMAYNE M. Blogging, citizenship, and the future of media[M]. Routledge, 2012.

[14] ANDROUTSELLIS-THEOTOKIS S. A survey of peer-to-peer file sharing technologies[J]. Athens University of Economics and Business, Greece, 2002: 1-31.

[15] 史建焘. P2P 文件共享系统安全性研究[D]. 哈尔滨: 哈尔滨工业大学, 2012.
SHI J T. Research on the security of P2P file sharing system[D]. Harbin, China: Harbin Institute of Technology, 2012.

[16] CLAY S. Listening to napster[J]. Peer-to-Peer: Hamessing the Benefits of A Disruptive Technology, 2001: 21-37.

[17] QIU D Y, RAYADURGAM S. Modeling and performance analysis of BitTorrent-like peer-to-peer network[J]. ACM SIGCOMM Computer Communication Review. ACM, 2004, 34(4): 367-378.

[18] MATEI R. Peer-to-peer architecture case study: gnutella network[C]//First International Conference on Peer-to-Peer Computing. IEEE, 2001: 99-100.

[19] YORAM K, DANNY B. The eMule protocol specification[EB/OL]. http://sourceforge.net, 2005.

[20] ZHU Z S, LU G H, CHEN Y, et al. Botnet research survey[C]//32nd Annual IEEE International Computer Software and Applications Conference. Turku, Finland, 2008: 967-972.

[21] 诸葛建伟, 韩心慧, 周勇林, 等. 僵尸网络研究与进展[J]. 软件学报, 2008, 19(3): 702-715.
ZHUGE J W, HAN X H, ZHOU Y L, et al. Research and development of botnets[J]. Journal of Software, 2008, 19(3): 702-715.

[22] GU G F, YEGNESWARAN V, PORRAS P, et al. Active botnet probing to identify obscure command and control channels[C]//Proceedings of 2009 Annual Computer Security Applications Conference. Honolulu, Hawaii, 2009: 241-253.

[23] DAGON D, GU G F, LEE C P, et al. A taxonomy of botnet structures[C]//The 23rd Annual Computer Security Applications Conference. Miami Beach, FL, USA, 2007: 325-339.

[24] STANLEY G N, AARON K. Usability and privacy: a study of Kazaa P2P file-sharing[C]//The SIGCHI Conference on Human Factors in Computing Systems. ACM, 2003: 137-144.

[25] LEDER F, WERNER T, MARTINI P. Proactive botnet countermeasures - an offensive approach[C]//1st CCDCoE Conference on Cyber Warfare. Tallinn, Estonia, 2009: 211-225.

[26] OLLMANN G. Botnet communication topologies[R]. Atlanta, GA: Damballa Inc, Technical Report: 2009-06-04, 2009.

[27] ROY F, JAMES G, JEFF M, et al. Hypertext transfer protocol--HTTP/1.1[R]. 1999.

[28] ZOU X G, LI Q, SUN S H, et al. The research on information hiding based on command sequence of FTP protocol[C]//Knowledge-Based Intelligent Information and Engineering Systems. Springer Berlin Heidelberg, 2005: 1079-1085.

[29] JARKKO O, DARREN R. Internet relay chat (irc) protocol[J]. IETF, Request for Comments (RFC), 1993, 1459.

[30] MARK D, JONATHAN R, HIROYASU S. A model for presence and instant messaging[R]. 2000.

[31] SINGH K, SRIVASTAVA A, GIFFIN J, et al. Evaluating email feasibility for botnet command and control[C]//38th Annual IEEE/IFIP International Conference on Dependable Systems and Networks. Anchorage, USA, 2008: 376-385.

[32] CHO C Y, CABALLERO J, GRIER C, et al. Insights from the inside: a view of botnet management from infiltration[C]//The 3rd USENIX Workshop on Large-Scale Exploits and Emergent Threats. San Jose, CA, USA, 2010: 2.

[33] SUN Y, LIU F M, LI B, et al. Fs2you: peer-assisted semi-persistent online storage at a large scale[C]//IEEE INFOCOM 2009. IEEE, 2009: 873-881.

[34] HECKMANN O, BOCK A, MAUTHE A, et al. The edonkey file-sharing network[J]. INFORMATIK, 2004, 51: 224-228.

[35] PORRAS P, SAIDI H, YEGNESWARAN V. An analysis of the iKee B iPhone botnet[J]. Lecture Notes of the Institute for Computer Sciences,

Social Informatics and Telecommunications Engineering, 2010, 47(5): 141-152.

[36] PETER U, MARTIN H, KAPITZA R, et al. Eliminating single points of failure in software-based redundancy[C]//2012 Ninth European Dependable Computing Conference (EDCC). IEEE, 2012: 49-60

[37] STOICA I, MORRIS R, KARGER D, et al. Chord: a scalable peer-to-peer lookup service for internet application[C]//ACM SIGCOMM 2001. New York: ACM, 2001: 149-160.

[38] RATNASAMY S, FRANCIS P, HANDLY M. A scalable content-addressable network[C]//ACM SIGCOMM 2001. San Diego: ACM Press, 2001: 161-172.

[39] ZHAO Y B, KUBIATOWICZ J, JOSEPH A D. Tapestry: an infrastructure for fault-tolerant wide-area location and routing CSD-01-1141[R]. California: University of California Berkley, 2001.

[40] MAYMOUNKOV P, MAZIERES D. Kademlia: a peer-to-peer information system based on the XOR metric[C]//International Workshop on Peer-to-Peer Systems 2002. Massachusetts: Springer Berlin, 2002: 53-65.

[41] SANDEEP S, ANDREAS T. Measuring the storm worm network[R]. HiNRG Technical Report: 01-10-2007, 2007.

[42] CLARKE I, SANDBERG O, WILEY B, et al. Freenet: a distributed anonymous information storage and retrieval system[C]//The Workshop on Design Issues in Anonymity and Unobservability. Berkeley, CA, USA, 2000: 311-320.

[43] DOUCEUR J R. The Sybil attack[J]. Peer-to-Peer Systems. Springer Berlin Heidelberg, 2002: 251-260.

[44] NUNNERY C, SINCLAIR G, KANG B B. Tumbling down the rabbit hole: exploring the idiosyncrasies of botmaster systems in a multi-tier botnet infrastructure[C]//The USENIX Workshop on Large-Scale Exploits and Emergent Threats. San Jose, CA, USA, 2010: 1.

[45] 蒋君. eMule 系统中的覆盖网络研究[D]. 上海: 上海交通大学, 2008.
JIANG J. The study of overlay network in eMule system[D]. Shanghai: Shanghai Jiao Tong University, 2008.

[46] TODD H, JOSE O, TONY M, et al. Active measurement data analysis techniques[EB/OL]. htpp://amp.nlanr.net, 2000.

[47] STEFAN S, KRISHNA G P, GRIBBLE S D. A measurement study of Napster and Gnutella as examples of peer-to-peer file sharing systems[J]. ACM SIGCOMM Computer Communication Review, 2002, 32(1): 82-82.

[48] RIPEANU M, FOSTER I. Mapping the gnutella network: macroscopic properties of large-scale peer-to-peer systems[J]. Peer-to-Peer Systems. Lecture Notes in Computer Science, 2002, 2429: 85-93.

[49] STUTZBACH D, REJAIE R. Characterizing the two-tier gnutella topology[C]//The 2005 ACM SIGMETRICS. 2005: 402-403.

[50] STEINER M, CARRA D, BIERSACK E W. Long term study of peer behavior in the KAD DHT[J]. IEEE/ACM Transactions on Networking, 2009.

[51] STEINER M, CARRA D, BIERSACK E W. Evaluating and improving the content access in KAD[J]. Peer-to-Peer Networking and Applications, 2010, 3(2): 115-128.

[52] 余杰. P2P 网络测量与安全关键技术研究[D]. 长沙: 国防科学技术大学, 2010.
YU J. Research on measurement and security of P2P networks[D]. Changsha: China National University of Defense Technology, 2010.

[53] NLANR M. Passive measurement and analysis[EB/OL]. http://prna.nlanr.net/PMA. 2003.

[54] HUANG L S, WANG W Y, LI C C, et al. Network fault analysis from passive measurement[J]. China Communications, 2012, 9(5): 64-74.

[55] CHRISTION N, WEIGEND A, CHUANG J. Content availability, pollution and poisoning in peer-to-peer file sharing networks[J]. Electronic Commerce, 2005: 1-10.

[56] BRUNNER R. A performance evaluation of the kad-protocol[D]. Mannheim, German: University of Manheim, 2006.

[57] STEINER M, ENNAJJARY T, BIERSACK E W. A global view of KAD[C]//Internet Measurement Conference (IMC). 2007.

[58] STEINER M, BIERSACK E W, ENNAJJARY T. Actively monitoring peers in KAD[C]//The 6th International Workshop on Peer-to-Peer Systems (IPTPS'07). 2007.

[59] HOLZ T, STEINER M, DAHL F, et al. Measurements and mitigation of peer-to-peer-based botnets: a case study on storm worm[C]//The First USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET'08). 2008.

[60] WANG T Z, WANG H M, LIU B, et al. Further analyzing the sybil attack in mitigating peer-to-peer botnets[J]. KSII Transactions on Internet & Information Systems, 2012, 6(10): 2731-2749.

[61] LU Q, LIU B, HU H P, et al. SMCSN: a new secure model of content sharing network by using multi-roles sybil nodes[C]//The fifth International conference on Computer Engineering and Networks. China, Shanghai, 2015.

[62] CARLTON R D, FERNANDEZ J M. NEVILLE S, et al. Sybil attacks as a mitigation strategy against the storm botnet[C]//3rd Internal Conference on Malicious and Unwanted Software. Alexandria, VA, USA, 2008: 32-40.

[63] CARLTON R D, FERNANDEZ J M, NEVILLE S. Optimising Sybil attacks against p2p-based botnets[C]//The 4th International Conference on Malicious and Unwanted Software. Montreal, Quebec, Canada, 2009: 78-87.

[64] 史建焘, 张宏莉, 方滨兴. BitTorrent 假块污染攻击的对抗方法研究[J]. 计算机学报, 2011, 34(1): 15-24.
SHI J T, ZHANG H L, FANG B X. Study on the countermeasures of bittorrent fake block attack[J]. Chinese Journal of Computers, 2011, 34(1): 15-24.

[65] KONG J, CAI W D, WANG L, et al. A study of pollution on BitTorrent[C]//The 2nd International Conference on Computer and Automation Engineering. Singapore, 2010: 118-122.

[66] SANTOS F R, CORDEIRO W L, GASPARY L P, et al. Choking polluters in bittorrent file sharing communities[C]//Network Operations and Management Symposium (NOMS) 2010. IEEE. Osaka, New Jersey, 2010: 559-566.

[67] KONG J, CAI W D, WANG L, et al. The evaluation of index poisoning in BitTorrent[C]// The Second International Conference on Com-

munication Software and Networks. IEEE. Singapore, 2010: 382-386.

[68] LOU X S, HWANG K. Collusive piracy prevention in P2P content delivery networks[J]. IEEE Transactions on Computers, 2009, 58(7): 970-983.

[69] LOCHER T, MYSICKA D, SCHMID S, et al. Poisoning the kad network[J]. Distributed Computing and Networking. Heidelberg, Springer. 2010: 195-206.

[70] ALTMANN J, BEDANE Z B. A P2P file sharing network topology formation algorithm based on social network information[C]//IEEE INFOCOM Workshops 2009. IEEE, 2009: 1-6.

[71] SHEN H Y, LI Z, CHEN K. Social-P2P: an online social network based P2P file sharing system[J]. IEEE Transactions on Parallel and Distributed Systems, 2015, 26(10): 2428-2440.

[72] MAHESWARA P V, NAIK K B. Routing protocol performance issues and evaluation considerations in MANET[J]. International Journal of Engineering Research and Applications (IJERA), 2013, 3(4): 1-5.

[73] ZHANG H, SHEN H. A social network based file sharing system in mobile peer to peer networks[C]//The 18th International Conference on Computer Communications and Networks. 2009:1-6.

[74] MAGALHAES J, HOLANDA M. EIKO: a social mobile network for MANET[C]//The Information Systems and Technologies. 2011: 8-15.

[75] FENG C S, YANG J, QIN Z, et al. Modeling and analysis of passive worm propagation in the P2P file-sharing network[J]. Simulation Modeling Practice and Theory, 2015, 51: 87-99.

[76] HE L, KYOUNGSOO B, JAESOO Y. A mobile social network for efficient contents sharing and searches[J]. Computers & Electrical Engineering, 2015, 41: 288-300.

[77] LIU G X, SHEN H Y, LEE W. An efficient and trustworthy P2P and social network integrated file sharing system[J]. IEEE Transactions on Computers, 2015, 64 (1): 54-70.

[78] 王天佐, 王怀民, 刘波, 等. 僵尸网络中的关键问题[J]. 计算机学报, 2012, 35(6): 1192-1208.

WANG T Z, WANG H M, LIU B, et al. Development of the research on some critical problems of botnets[J]. Chinese Journal of Computers, 2012, 35(6): 1192-1208.

作者简介：

鲁强（1987-），男，湖北随州人，国防科学技术大学博士生，主要研究方向为网络与信息安全。

刘波（1973-），男，湖北仙桃人，博士，国防科学技术大学研究员，主要研究方向为网络与信息安全。

胡华平（1967-），男，江西临川人，博士，国防科学技术大学研究员，主要研究方向为网络与信息安全。