信息通信领域产学研合作特色期刊
第三届国家期刊奖百种重点期刊 | 中国科技核心期刊

ISSN 1009-6868
CN 34-1228/TN

中兴通讯技术

ZTE TECHNOLOGY JOURNAL

www.zte.com.cn/magazine

2016年2月 • 第1期

专题：网络空间安全

9 771009 686168 02>

中兴通讯技术
ZTE TECHNOLOGY JOURNAL
二〇一六年二月 第二十二卷第一期 总第一二六期

《中兴通讯技术》第7届编辑委员会委员名单

信息通信领域产学研合作特色期刊
第三届国家期刊奖百种重点期刊
中国科技核心期刊
工信部优秀科技期刊
中国五大文献数据库收录期刊
ISSN 1009-6868
CN 34-1228/TN
1995年创刊

办刊宗旨

以人为本，荟萃通信技术领域精英；
迎接挑战，把握世界通信技术动态；
立即行动，求解通信发展疑难课题；
励精图治，促进民族信息产业崛起。

Contents 目次

中兴通讯技术 总第126期 第22卷 第1期 2016年2月

卷首特稿

专题：网络空间安全

专家论坛

企业视界

技术广角

综合信息

期刊基本参数：CN 34-1228/TN*1995*b*16*64*zh*P* ¥ 20.00*15000*15*2016-02

Contents 目次

ZTE TECHNOLOGY JOURNAL Vol. 22 No. 1 Feb. 2016

Guest Paper

Special Topic: Network Space Security

Expert Forum

Enterprise View

Technology Perspective

敬告读者

2016年第1—6期专题

1 网络空间安全
杨义先 北京邮电大学 教授
杨　庚 南京邮电大学 教授

2 大数据分析处理与应用
郑纬民 清华大学 教授

3 5G技术与业务创新
王　京 清华大学 教授
向际鹰 中兴通讯股份有限公司 博士

4 天地一体化信息网络
张乃通 中国工程院 院士
顾学迈 哈尔滨工业大学 教授

5 工业互联网与智慧工厂技术
邬贺铨 中国工程院 院士
王耀南 湖南大学 教授

6 SDN/NFV的实践与规模应用
蒋林涛 中国信息通信研究院 教授

专题栏目策划人

杨义先

北京邮电大学教授、博士生导师、信息安全中心主任，灾备技术国家工程实验室主任，首批长江学者特聘教授，首届国家杰出青年基金获得者，中国密码学会副理事长；目前研究方向为网络空间安全、现代密码学和纠错编码等；获得包括国家发明奖和省部级科技进步奖等在内的各类科技奖励20余项，主持和参与多项国家“863”计划项目、国家自然科学基金项目、省部级等科研项目；发表高水平论文500余篇，出版专著及教材20多部，持有发明专利4项。

杨 庚

南京邮电大学研究生院常务副院长，中国计算机学会、中国通信学会高级会员，教育部高等学校信息安全专业教学指导委员会委员、《计算机通信与网络》国家精品课程负责人和国家级精品视频共享课负责人；目前研究方向为云计算与大数据安全、信息与网络安全、分布与并行计算等；曾获省部级科技进步奖2项，主持和参与国家“973”项目课题、国家“863”计划项目、国家自然基金项目等8项，主持省级科研项目10项；近年来发表论文近百篇，持有中国发明专利20余项，已转让4项。

专家论坛栏目策划人

李 军

清华大学信息技术研究院院长，清华信息科学与技术国家实验室常务副主任，中国电子学会计算机工程与应用分会副主任，工信部电子科技委委员；目前研究方向为网络信息分类与过滤、网络流量观测与控制、数据中心网络虚拟化等算法和系统方面的研究；作为“863”目标导向课题“一体化网络数据深度安全检测与分析的技术与系统”项目负责人，获得2014年中国电子学会科学技术奖二等奖；著译中外教材3部，发表学术论文百余篇，持有美国专利2项，中国发明专利15项。

专题导读

2015年，对中国信息安全教研界来说，最大的新闻莫过于“网络空间安全”被国务院学位委员会正式批准为国家一级学科。从此，高校的“网络空间安全学院”将如雨后春笋般大量出现，信息安全高级人才的培养将走上快车道。在这一关键时刻，如果没有做好网络空间安全体系的设计，积极探讨网络空间安全的关键技术，那么很有可能会事倍功半。本专题的目的之一，就是希望借助《中兴通讯技术》，促进中国网络空间安全界尽可能多的同仁们形成合力，争取事半功倍！

网络空间安全是一个系统工程，必须重视顶层设计，必须从体系架构方面给予全面考虑。《网络空间安全面临的挑战及应对策略》和《网络空间安全体系及关键技术》两篇文章，在分析了中国网络空间安全现状及其面临的威胁后，探讨了应对策略；并通过分析传统线性结构防御体系以及传统的网络空间安全问题，提出了新型立体式网络空间安全体系结构，指明了网络空间安全技术的发展趋势和未来的研究方向。

“大数据安全”是网络空间安全的难点和重点，也是未来的热点。《云时代下的大数据安全技术》和《面向数据的安全体系结构初步研究》两篇文章，分别从不同的角度，研究了大数据面临的威胁与挑战，从大数据安全体系、Hadoop安全架构、数据所有权确立、数据注册、防范APT攻击技术等方面提出解决方案，同时也提出了以数据为核心和面向数据的信息安全解决方案，即面向数据的安全体系结构（DOSA）。

云计算、大数据、物联网是“互联网+”时代关键的技术，然而如果没有安全计算，势必会严重影响它的发展。《安全多方计算技术研究与应用》一文，从介绍安全多方计算的概念开始，基于同态密码、格理论密码设计了实用的安全多方计算协议，并总结展望了安全多方计算协议研究发展趋势。

网络空间的命脉是安全，安全的核心是密码，而同态加密又是密码研究的最新热点。《同态加密的发展及应用》一文，综述了同态加密方案的发展，介绍了同态加密在安全多方计算、密文检索、安全云计算、电子选举等诸多方面的应用。还着重分析了一些部分同态、浅同态和全同态加密方案的优缺点，介绍了同态加密中存在的一些公开问题。

在数字时代，人们的生活越来越依靠于各种数据，“好事”和“坏事”都会在数字空间中进行，因此，“数字取证”就成为了网络空间安全的一个重要方面。数字取证也可称为计算机法医学，它是指把数字空间看做犯罪现场，运用先进的辨析技术，对电脑犯罪行为进行法医式的解剖，搜寻确认罪犯及其犯罪证据，并据此提起诉讼。《计算机网络取证和调查的科学研究》一文，通过一个真实的网络犯罪案例的犯罪现场重建，阐述了调查取证基本原则。并在此基础上，通过理论和实验分析，将取证科学过程应用到对P2P网络中某些文件的首次上传者的调查中。

这期专题主要和大家一起来讨论网络空间的关键、热门技术。这些论文凝聚了作者多年的研究成果和工作经验，希望能给读者有益的启示与参考。在此，对各位作者的积极支持和辛勤工作表示衷心的感谢！

杨义先　杨 庚

2015年12月21日

DOI:10.3969/j.issn.1009-6868.2016.01.001
网络出版地址:http://www.cnki.net/kcms/detail/34.1228.TN.20151217.1002.002.html

[摘要] 通过分析ICT发展对中国经济和社会可持续发展的重要意义,指出在中国发展ICT,重点在于发展高性能核心电子及光子器件,并提出了目前两项最具影响力的新型技术——三维集成技术和光电子技术。认为三维集成技术以其独特优势,成为未来微纳光电子的发展方向;而光电子技术在ICT发展中也正扮演着越来越重要的角色,是未来超高速通信的桥梁。

[关键词] ICT;核心电子器件;摩尔定律;三维集成技术;光电子技术

[Abstract] Information communication technology (ICT) plays a significant role in the sustainable development of both the economy and society of China. The focus of ICT in China is development of core electronic and opto-electronic components. Two of the most impressive emerging technologies are 3D integration technology and opto-electronic technology. 3D integration technology is the future development direction of micro/nano optoelectronic technology in the future. The development of the electronic technology in the ICT is also playing a more and more important role, and which becomes the bridge of future ultra high-speed communications.

[Key words] ICT; core electronic components; Moore law; 3D integration technology; opto-electronic technology

中图分类号:TN929.5 文献标识码:A 文章编号:1009-6868 (2016) 01-0002-03

ICT核心器件发展展望

Development Prospects of ICT Core Components

李尔平/LI Erping

(浙江大学 信息学部,浙江 杭州 310027)
(Faculty of Information Technology, Zhejiang University, Hangzhou 310027, China)

李尔平,浙江大学信息学部副主任、教育部“长江学者”讲座教授、IEEE Fellow、首批国家千人计划特聘教授;研究领域为新型微纳光电子器件集成技术、微波电子射频天线及电磁兼容等;荣获多个国际奖项和荣誉,包括2015年荣获国际IEEE最高奖之一——理查德-司徒达特奖;发表国际论文400余篇,英文专著两部,申报多项美国、新加坡和中国专利。

收稿日期:2015-11-25
网络出版时间:2015-12-17

信息通信产业是支撑中国经济与社会发展的关键支柱产业之一,尤其是进入21世纪以来,世界经济的信息化和全球化发展趋势日益明显,ICT促进中国整体产业结构调整、转换和升级,成为推动中国经济增长的重要手段,支撑着国家的可持续发展。

大力发展ICT是中国中长期科学和技术发展方向的重要课题,符合国家的重大战略需求。在《国家中长期科学和技术发展规划纲要(2006—2020年)》中提到“重点开发高性能的核心网络设备和传输设备、接入设备,建立可信的网络管理体系,开发智能终端和家庭网络等设备和系统,支持多媒体、网络计算等宽带、安全、泛在的多种新业务与应用[1]”。2015年5月,国务院公布《中国制造2025》规划,在新一代ICT产业中阐述:“掌握新型计算、高速互联、先进存储、体系化安全保障等核心技术,超高速大用量智能光传输技术、未来网络核心技术和体系架构,推动核心信息通信设备体系化发展与规模化应用[2]”。

经过30多年的发展,中国信息通信系统整机和设备制造能力已经位居世界前列,例如在光通信领域,自主整机产品已占全球网络的2/5,光接入设备占全球的3/4,光纤光缆占全球的1/2。但中国在ICT核心电子器件技术领域还处于跟跑阶段,单在2013年中国集成电路进口2 322亿美元,超过原油进口,信息产业的“芯”在外。制约中国ICT核心电子技术发展的突出问题包括电子器件加工设备研发实力薄弱,缺乏标准化和规范化的电子器件工艺平台以及芯片模块化封装和测试分析技术落后等。ICT核心电子器件在无线通信、高性能计算、智能交通、远程医疗、航空航天和深空探测,以及与国家安全相关的空天平台上通信、雷达、导航和控制系统中的应用举足轻重。因

此，为了确保中国ICT在国际上的可持续引领作用和对国民经济的重大贡献，必须从基础的核心电子器件出发，重点发展ICT核心电子器件的相关技术。

1 ICT核心器件发展的重大技术挑战

半个世纪以来，微纳光电子技术一直遵循摩尔定律，即芯片的集成度每18个月翻一番。对于电子器件小型化和多功能的需求已促使集成电路内的晶体管尺寸减小到了10 nm以下。但是，这一缩小的趋势受到了严峻挑战，其关键问题在于几下几点。

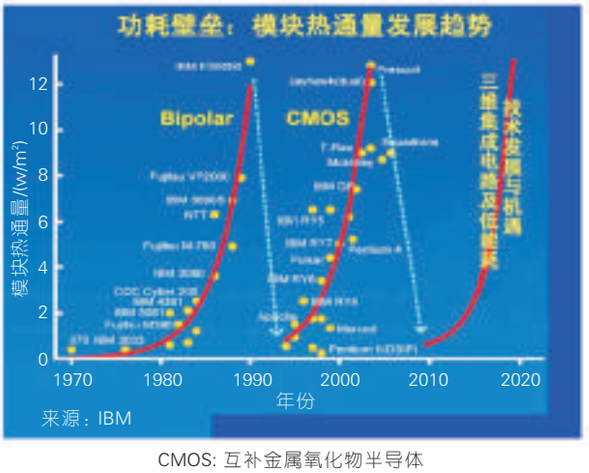
（1）互补金属氧化物半导体（CMOS）工艺已经逼近物理极限。首先，随着特征尺寸的不断降低，金属互连的延迟、功耗和噪声等在不断增加，互连已经取代晶体管成为决定集成电路性能的主要因素及限制其未来发展的真正瓶颈。其次，能耗密度的急剧增加也限制了高性能集成电路的发展。高度集成电路的能耗密度大幅提高，已经达到每平方厘米100 W，使得芯片直接散热和冷却技术面临很大的挑战。如图1所示，多核高性能处理器对数据带宽的要求不断增加，而现有的互连以及目前的二维平面集成技术已经无法满足多核处理器对高速、宽带的数据通信的要求，功耗问题已是影响时钟速率提升的关键问题。最后，更小尺寸的硅加工技术也是微电子进一步小型化面临的挑战。

（2）系统集成各种功能模块需要多种工艺的支持。当前的嵌入式系统依赖系统芯片（SoC）实现，SoC集成可实现存储器、数字电路、模拟电路、射频电路、大功率电路、光/电器件、电源管理电路、传感器电路等多种功能模块的集成。有些电路模块在特征尺寸减小时并不一定能提高性能，如射频电路模块。由于以上原因，很难将众多的电路均以最佳性能、利用同一工艺集成到一块SoC芯片中，或者单片SoC集成的成本过高。此外，高密度集成导致管脚开销急剧增加。

（3）现有光电器件的系统集成方式速度低、损耗大。无线基站系统通常需要由光纤传输的光信号和由无线信道传输的射频信号之间进行转换。由于光信号和电信号的处理电路依赖不同工艺，无法在单一芯片上集成，因此在现有的系统方案中，这两部分电路只能以分立器件进行系统集成，这种方式造成信号转换过程中的损耗极大而且速度低，大大降低了系统的性能。

（4）系统级封装（SiP）是一种被广泛采用的高集成度的方法，通过直接在封装层次集成多个芯片来提高集成度并解决不同工艺集成的问题。但是一些SiP电路中需要用小尺寸工艺芯片驱动较大的外接负载，功耗依然很高。

为解决如上所述的问题，业界提出了几个重要的技术，其中最具影响力的为垂直互连的三维器件集成技术和光电子技术。



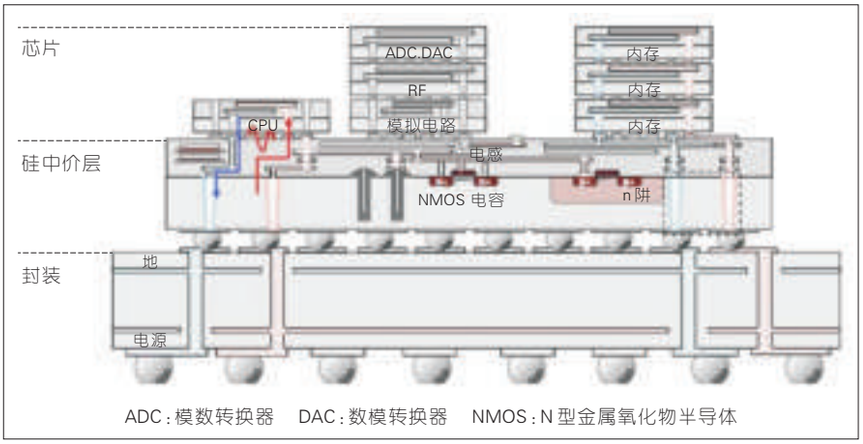
CMOS：互补金属氧化物半导体

◀图1
半导体技术的功耗发展及其技术预测

2 三维集成技术——超越摩尔定律

如图2所示的三维集成技术的基本思想是在垂直方向叠加芯片，并使用垂直硅通孔（TSV）互连[3]为各芯片之间提供直接、短距离的电路连接。三维集成电子器件和硅通孔技术使芯片能够更有效地安排空间版图，减少连线长度，总体连接缩短大约30%～40%，从而显著提高封装密度，提高芯片速度。此外，三维集成允许不同工艺的芯片以自然的方式集成，如可以将无线芯片、传感器、光子器件、微机电系统（MEMS）及CMOS芯片集成在一个芯片上。因此，三维集成技术为目前集成电路面临的数据传输带宽、芯片功耗和速度以及异质芯片集成等问题提供了切实可行的解决途径[4]。三维集成电路和硅通孔技术将推动半导体行业向延续摩尔定律和超越摩尔定律发展，国际半导体技术发展蓝图（ITRS）也将以垂直硅通孔为基础的三维集成列为微纳光电子科学与技术发展的一个非常重要的方向[5]。

三维集成技术对集成电路领域产生了巨大的影响，也带来了一系列挑战。首先，由于片上有许多硅通孔（几千甚至上万），需要占用片上面积，必须用小孔径以减少通孔所占用面积，因此研究小孔径及高深宽比的制造工艺及集成技术，是极具挑战的技术。其次，三维集成的散热问题比二维集成重要许多，这是由于三维高集成度使电路系统的“表面积-体积”比下降很多，进而使得散热更加困难，寻求更先进热管理技术对于三维集成至关重要。除此之外，高度集成



▲图2 垂直硅通孔互连的三维集成器件

的电路结构对于电、热、应力的可靠性研究，设计方法、设计规则以及测量测试方法也都提出了新的要求。综上所述，三维集成有其独特的优势，是未来微纳光电子的发展方向，但目前还存在着关键技术挑战，如热管理问题、垂直信号传输机制与制造技术等问题，解决这些关键技术是未来的研究重点。

3 光电子技术——实现超高速通信

光电子技术在ICT中正扮演着越来越重要的角色[6]。在宽带通信网络方面，统计数据显示：传输和交换设备中光器件的成本比重高达70%和60%，对于高端的网络通信设备这个比例更高。而在移动通信方面，随着接入带宽需求的快速发展（从3G的10 Mbit/s，到4G的100 Mbit/s，甚至未来5G的1 Gbit/s），移动的灵活性必须与光纤的宽带相结合，光与无线融合已经是大势所趋，光电子技术将成为移动通信繁荣4G，推广5G，甚至部署后5G（B5G）技术的重要支撑。此外，在高性能计算方面，基于光电子器件的柜间光互联（即计算机机柜之间）已经走向板间光互联（即单台计算机电路板之间），未来更将走向片上光互联（即芯片内部之间），才能够支撑超级计算机的性能进一步提升。图3为Intel报道的片间光子集成与IBM报道的片上光子集成的发展趋势。

▲图3 Intel报道的片间光子集成与IBM报道的片上光子集成

尽管光电子技术在近几年里发展迅速，但是由于总体发展时间短，该技术在未来较长时间内仍会存在诸多挑战。首先，集成化是光电子技术发展的必由之路，这是光电子科学技术发展已经证明了的事实，也是学界和产业界的共识。只有集成才能够支撑未来信息系统对速率宽带、能耗体积以及智能可调控等发展需求。然而，光电子集成与基于单一硅材料的微电子集成有着巨大的差别，涉及异质、异构、异速、异维等关键科学问题。其次，随着各种信息应用的动态化和复杂化，为了支撑信息传输、交换及处理的智能化需求，光电子功能模块必须具备可重构的能力，并且能够在对系统性能没有影响的情况下完成所需要的快速功能转换。最后，为了适应现代信息社会对宽带需求的高速增长，核心光网络系统所面临的信号维度已经不是传统的单一维度（如波长波分复用（WDM）、幅度二进制启闭键控（OOK）等），而是越来越复杂的维度组合（波长、偏振、模式、轨道角动量等等），因此提出高效的信号调控机制（包括产生、传输、交换、接收等）已经成为日益紧迫的要求。总的来说，光电子技术是未来超高速通信的桥梁，而芯片与器件层面的光电子集成，功能模块层面的可重构智能化，以及系统应用层面的多维信号调控将成为下一阶段的研究重点。

参考文献
[1] 国家中长期科学和技术发展规划纲要（2006–2020）[EB/OL]. http://www.most.gov.cn/kjgh/kjghzcq/
[2] 国务院印发《中国制造2025》[EB/OL]. http://www.gov.cn/xinwen/2015–05/19/content_2864538.htm
[3] LI E P. Electrical Modeling and Design for 3D System Integration [M]. USA: John Wiley & Sons, Inc., 2012
[4] 王喆垚. 三维集成技术[M]. 北京:清华大学出版社, 2014
[5] 国际半导体技术发展蓝图[EB/OL]. http://www.itrs.net/
[6] LI E P, CHU H S. Plasmonics Nanoelectronic and Sensing Devices [M]. UK: Cambridge University Press, 2014

DOI:10.3969/j.issn.1009-6868.2016.01.002
网络出版地址:http://www.cnki.net/kcms/detail/34.1228.TN.20151117.1506.006.html

网络空间安全面临的挑战及应对策略

Challenges and Countermeasures of Network Space Security

周延森/ZHOU Yansen
周琳娜/ZHOU Linna
(国际关系学院 信息科技学院,北京 100091)
(Department of Information Science and Technology, University of International Relations, Beijing 100091, China)

中图分类号:TP393.4 文献标志码:A 文章编号:1009-6868 (2016) 01-0005-005

摘要: 当前网络空间安全威胁包括:大规模分布式拒绝服务攻击,众多主机被境外结构控制,关键信息存储在非本国品牌服务器中,用户借助翻墙软件逃过监控等。提出了应对网络空间威胁的措施——自主可控与互联网应用创新,具体包括自主研发CPU和路由器等核心硬件以及操作系统和数据库等系统软件。认为下一代信息技术从内容的深度和广度挑战现存的网络空间,形成了新的安全威胁,这需要国家加快下一代网络的部署及研发。

关键词: 网络空间;空间主权;大数据;物联网;云计算

Abstract: The main security threats of network space include large scale distributed denial of service attack denial of service attacks; many hosts controlled by foreign institutions; key information stored in many foreign brand servers; and users with special software who can escape monitoring. The main measures to deal with space threats are self controlled and Internet application innovation, including 1) independent research and development of core hardware including CPU and router and 2) system software including operating systems and databases. The next generation of information technology from the depth and breadth of content challenges the existing network cyber, which forms a new security threat. This requires China to speed up the research and deployment of the next generation network.

Key words: network space; spatial sovereignty; big data; Internet of things; cloud computing

随着互联网应用不断深入,网络空间逐渐被视为继陆、海、空、天之后的"第5空间",成为世界关注的焦点和热点。网络在方便和丰富人们生活的同时,也使得网络攻击行为广泛存在。互联网的应用已深入到社会的方方面面,小到百姓的日常生活,例如网上购物和网络金融,大到涉及国计民生的行业,例如国家电网等。西方国家已将网络空间安全提升到国家战略高度予以重视,中国政府也非常重视网络空间安全。国家教育部最近将网络空间安全专业列为一级学科,这说明教育部门正在为国家网络空间安全提供强大可持续的人才储备。

从国家层面上看,为了适应网络安全空间面临的严重挑战和维护国家的网络主权安全,2014年中央网络安全和信息化领导小组正式成立,这为网络空间安全提供了强大的组织保证。习近平总书记说:"没有网络空间安全就没有国家安全"。这充分说明党和国家领导人高度重视网络空间安全,把网络空间安全提升到国家主权安全的高度[1]。

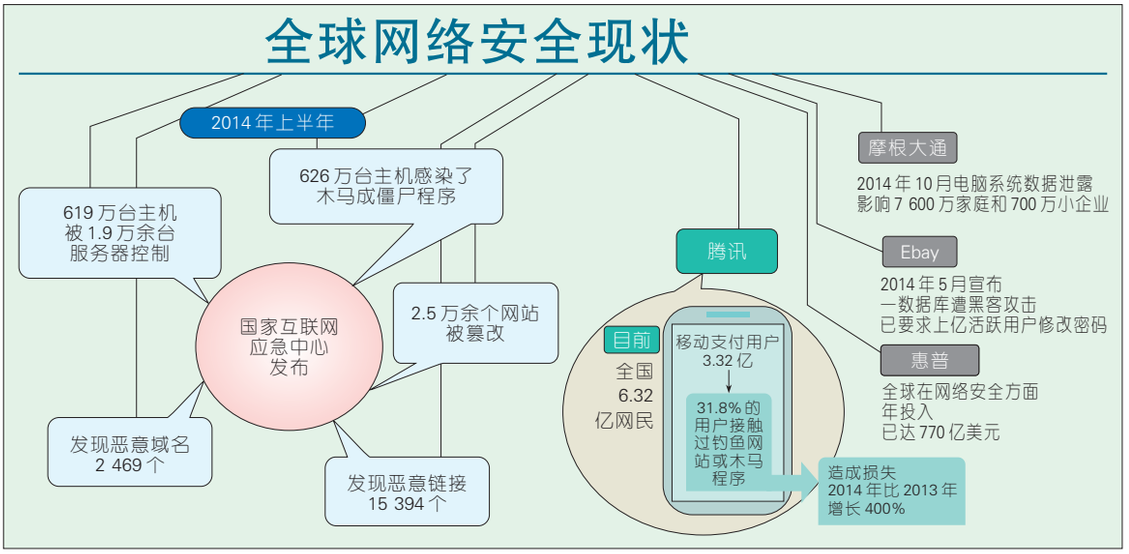
网络空间不是一个抽象虚拟的概念,而是实实在在存在于我们的生活当中。从物理结构上分析,网络空间主要由通信基础实施、部署在网络外围的计算机、移动智能终端以及各种提供共享资源的服务器等组成。此外,各种互联网应用和提供的服务都是网络空间重要的组成部分。

当前,中国网络空间安全形势非常严峻。图1列出了2014年上半年中国网络空间安全现状的一些数据[2]。

1 网络空间安全面临的挑战

随着信息技术的迅猛发展和互联网的普及,特别是以微信、Facebook和LINE为代表的新一代即时通信软件的推广和普及应用,使得信息传播的速度、广度和实时性都达到史无前例。互联网应用正在深入到国家与社会的各个方面,同时也伴随着大量的不良信息以及恶意的网络行为,如计算机木马、拒绝服务攻击、垃圾邮件、恐怖主义视频以及泄露的党和国家机密信息等。网络不良信息和网络恶意行为不仅会造成重大的经济损失,而且会严重威胁国家的政治、经济、国防、文化等正常秩序,干扰人民群众的正常生活,甚至会因为恶意散发的网络谣言会引发国家与社会

收稿日期:2015-11-15
网络出版时间:2015-11-17
基金项目:自然科学基金(61170175)



◀图1
网络空间安全现状

动荡。由此可见，网络空间安全在国家主权安全中的地位和作用也越来越大[3]。

互联网已经成为党政机关和企事业单位获取信息的重要来源，来自于互联网的海量数据正成为各行各业做出正确决策的重要依据。因此在下一代网络空间中，针对海量数据基于数据挖掘算法和大数据处理技术以及云计算技术进行深度及时分析处理，是任何国家政府和企事业单位必须解决的关键问题。对大数据获取与智能处理是决定国家和社会发展的一个关键问题，是走出“有数据但无知识”困难局面的一个重要的突破口[4]。

（1）大规模分布式拒绝服务攻击

下一代网络IPv6只在网络层进行了深层次的改变[5]，传输层和应用层没有改变。因此，基于传输层的分布式拒绝服务攻击还会存在。据相关结构报告：中国大型的电子商务网站和提供搜索引擎服务的网站屡次遭到大规模的拒绝服务攻击（DDOS），例如天猫和京东等电网站都曾遭受过此类攻击，给网站的卖家带来巨大的经济损失；百度网站在过去几年也遭受此类攻击，由此导致众多用户无法使用搜索服务。

（2）众多主机和服务器受控制

中国众多的政府机关、企事业单位以及个人的服务器和主机遭到境外机构的控制。据国家互联网应急中心发布的数据显示：2014年，中国有650多万台主机因感染了木马病毒而被境外服务器控制，2万多个网站被篡改主页。这些被恶意控制的服务器和主机不仅泄漏了国家机密和私人隐私，而且成为境外机构发动分布式决绝服务攻击的帮凶[6]。

从这些攻击的源头上分析，我们发现大部分参与实际攻击的是这些被控主机，而实际上真正发起攻击的源头在其他一些国家。

（3）翻墙软件挑战信息监控

中国非常重视互联网内容的监管工作[7]，在国家多个关键的互联网接口处都有网络监控设施。但是，随着互联网翻墙软件的轻易获取，只要稍微懂得互联网技术的用户，就可以通过在计算机和手机上安装此类软件，突破网络监控，随意访问一些严禁访问的网站。例如，法轮功组织开发的无界限翻墙软件，能够在本地机器实现代理服务器的功能，通过该代理可避开监管，访问到大智慧等非法网站。另外，通过浏览器设置代理服务器的连接方式，能够完成网络攻击源IP地址的隐藏。因为在被攻击服务器的日志中总是留下代理服务器IP地址作为网络攻击行为源IP地址，除非各个国家政府之间进行互联网安全的协作，否则这种攻击行为难易完成源头的定位。

（4）骨干网的高速交换机和路由器受控制

在互联网技术方面，美国遥遥领先于其他国家[8]，特别是在骨干网的基础通信设施方面。目前，中国骨干网的大部分通信设备主要由美国思科等公司提供。路由器配有独立IP地址，路由器人为设置的漏洞等，均会导致这些生产路由器的公司远程控制这些路由器。这些公司完全可以做到通过远程切断中国网络的通信服务。

（5）网络空间海量数据处理面临挑战

互联网空间中的大数据包含结构化和非结构化两种类型[9]。据互联网相关结构调查发现：在网络海量数据中，约20%的数据是结构化的，约80%是非结构化或半结构化的，并且非结构化信息增长率是结构化信息增长率的两倍。对非结构化数据的处理需要大数据处理核心技术，而目前这些技术掌握在西方国家手中。海量数据用现行的检测过滤技术无法做到实时检测，因此很多入侵者就可以利用这个漏洞进行浑水摸鱼式

的系列攻击。

目前网络上的加密软件体现出加密算法强度大、密码长度超长的特征，暴力破解面临严重挑战。

（6）基础设施面临瘫痪

随着社会重要基础设施的高度信息化，社会的“命脉”和核心控制系统有可能面临损坏和瘫痪。这主要有两点原因：

• 关乎国计民生的海量数据存储在非本国品牌的数据库服务器中。数据库系统软件主要有微软的SQL Server、甲骨文的ORACLE以及IBM的DB2以及SYBACE。目前，中国的主要银行、铁路、民航、社保和其他关乎国计民生的行业都离不开上述的数据库系统的支持。由于这些软件都不是开源的，因此我们对其存在的安全漏洞无从所知。这些数据库系统软件公司可以配合所在国政府轻易地删除或修改中国的核心数据，而这些数据的丢失则会引起社会的恐慌，从而达到“不战而屈人之兵”的效果。

• 涉及国计民生的大型服务器都是西方国家的品牌。美国的IBM、惠普等公司占据中国大型服务器80%以上的市场，目前涉及的主要行业为金融、交通和电力等。这些服务器具有强大的计算能力，每天都在为政府机关和企事业单位以及个人提供各种各样的金融和电子商务等日常服务。如果这些服务器受到远程控制的话，可以随时被切断服务。这将会导致整个社会的混乱。

（7）根域名服务器受控制

根服务器主要用来管理互联网的主目录。所有的根服务器均由美国政府授权的互联网域名与号码分配机构——ICANN统一管理。ICANN还负责全球互联网域名根服务器、域名体系和IP地址等的管理。

这些逻辑根服务器可以指挥Internet Explorer这样的Web浏览器和电子邮件程序控制互联网通信。自互联网成立以来，世界对根域名服务器的依赖性非常大，美国通过控制根服务器而控制了整个互联网。从理论上说，任何形式的标准域名要想被解析，按照技术流程，都必须经过全球“层级式”域名解析体系的工作，才能完成。层级式域名解析体系第1层就是根服务器，它负责管理世界各国的域名信息；在根服务器下面是顶级域名服务器，即相关国家域名管理机构的数据库，如中国的CNNIC。美国可以通过其控制的根域名服务器，随时切断中国对外的互联网服务。

（8）计算机和智能手机的核心部件受控制

CPU是计算机和智能手机的核心部件和主要计算单元。在PC端，CPU主要有Intel和AMD两大品牌；智能手机所使用的CPU都由高通公司提供。如果在这些CPU上植入木马，现存的任何检测软件都无法检测到此类硬件木马和病毒，并且它们可根据需要随时引发病毒，造成计算机系统停止工作。

操作系统是软件之母，所有其他软件的运行必须依附在操作系统正常工作的基础上。在PC端，目前主要品牌有Windows操作系统；手机操作系统主有IOS和Android。此外，现有的信息安全软件无法检测到存在于CPU之上的硬件木马，也无法检测到存在于操作系统核心模块的软件木马，但其他国家研制这些软硬件的公司可以轻易远程控制这些木马。

2 网络空间安全挑战的应对策略

针对上述网络空间安全的一系列挑战，我们提出了应对策略——自主可控与创新。具体包括以下几个方面。

（1）自主研制CPU

到目前为止，世界上90%以上的CPU芯片都由美国的Intel、AMD和高通公司控制。为了摆脱这方面的被动局面，中国现有的研发与生产CPU的公司需要进行整合，争取在未来5年设计出1～2款具有国际竞争力的CPU，特别要为军队和党政机关研发自主可控的CPU。中国应该在未来十年，投入更多的资金加强这方面的研究和设计，以摆脱有机无芯的被动局面。

（2）自主研发操作系统

操作系统是计算机与智能手机的大脑，它们能够指挥多个进程协同和并发工作。经过中国计算机科学家多年的努力，在基于开源操作系统Linux的基础上，先后研发出红旗、银河麒麟和中国操作系统（COS）等，但目前这些系统在应用的广度和深度方面还有很多欠缺。中国应该制订操作系统研发的中长期国家计划，加大在操作系统研发方面的投入，特别是在手机操作系统方面。

（3）自主研制可控的服务器和大型数据库系统软件

服务器与国民经济息息相关[10]，关乎国计民生的数据都存放在大型数据库系统当中。因此，中国需要在未来十年，通过市场化等手段，加快国外知名品牌在高速服务器的核心技术转让。另外，还应该加强数据库等系统软件的研发和投入，组织1～2个核心科研团队，力争研发出具有市场竞争力的数据库系统。中国应该继续支持以山东浪潮为代表的生产大型服务器的高科技公司，在资金、市场等方面给予倾斜。

（4）使用国产品牌的路由器和交换机组织骨干网

目前，中国骨干网上的路由器和交换器主要由思科公司提供[11]。但是，随着中国国内以中兴通讯为代表的通信设备公司的崛起，中国应下定决心用国产的先进品牌逐渐替换非国产的品牌。为此，中国应该加大对这些产业的投入，在市场上加以引导，让大型的国有通信企业在采购设备中加大国产品牌的比重。

（5）加强互联网内容的管理以及控制

目前，以百度、阿里和腾讯为代

表的互联网公司在搜索引擎、电子商务和即时通信等新兴网络应用方面走在世界的前列。在带来可观的经济收入的同时也将大量的互联网数据留在了中国的服务器当中，减少了数据流量进入其他国家导致可能泄密的情况发生[12]。因此，中国需要通过扶持互联网新型应用，争取出现更多的大型互联网公司。

（6）大力支持目前中国的网络安全公司

中国网民的技术涵养近几年得到了很大的提高[13]，但是和西方发达国家相比，在网络安全意识方面还有一定的差距。另外，过去许多杀毒软件公司在升级病毒库时需要收取一定的服务费，许多用户不安装杀毒软件。上述这些因素都可能导致多达千万的中国的主机被其他国家机构与组织控制。以360为代表的新型网络安全公司，提供了一种全新的、免费的网络安全服务，使得中国大部分主机都安装了360的杀毒和防护软件，这些安全软件能够为用户提供绝大部分的安全服务，避免了更多的主机成为肉鸡。因此，中国应该加大以360为代表的网络安全公司的支持力度。

（7）部署根域名服务器于国际组织中管理

除了部分根域名服务器在欧洲与日本之外，大部分根域名服务器主要受控于美国商务部。由于美国参众两院的阻拦，美国迟迟无法将域名服务器转交给联合国等国际组织。随着信息技术实力的增强，广大发展中国家国家应该联合起来，要求美国交出根域名服务器的管理权限，将众多的域名服务器部署在一些发展中国家。中国应该在这个事件中起主导作用。

（8）建立国家级网络空间攻防专业技术队伍

中央网络安全领导小组的成立为网络空间的安全提供了组织上的保证，同时也还需要从技术上为网络空间的安全提供保障。中国应该整合国内所有研究网络空间安全的企事业单位以及科研院所，从中抽调精英，组建具有中国特色的网络安全部队，加强互联网翻墙软件的破译工作，对其他国家的代理服务器进行有效拦截。

3 下一代信息技术对网络空间安全的影响和应对措施

以下一代网络、物联网、大数据处理以及云计算为代表的新一代信息技术的最大特点就是促进了海量数据在更大范围的交流和超范围的信息共享[14]。新一代信息技术在给广大用户带来便利的同时，也给中国维护网络空间安全带来全新的挑战。

3.1 下一代信息技术对网络安全的影响

（1）下一代网络技术

下一代网络超快的网速为信息的交互带来极大的方便，但加密机制也给网络侦控带来了深远的影响。过去，国家机关可以借助网络监听技术，从不法分子的网络通信中获取犯罪证据；现在，对下一代网络存在的海量信息的监测，肯定会漏掉部分重要的敏感信息。由于通信协议中有强大的加密功能，大量的通信信息采用加密的方式，因此采用常规的监听和破解手段无法实现对加密信息有效的破解。如果敌对分子利用通信协议中的加密技术而导致信息加密无法破解，这就会给国家安全带来一定的危害。

（2）云计算

云计算是未来信息技术的一个重要发展方向。目前，云计算核心技术主要由美国几大IT巨头控制，包括核心的数据处理和海量数据的存储技术[15]。

目前，中国在云计算中心的建设发展没有国家统一的规划，发展有些混乱，如地方政府纷纷建设云计算平台，但因缺乏大数据处理的核心技术支持，导致云计算应用能力不够，缺乏业务数据。因此，中国应该制订科学合理的规划，对全国的云计算平台进行资源整合，构建具有自主核心技术、大规模的和具有核心竞争力的云计算平台，确保国家的核心数据能够留在本国的云计算中心进行处理。

另外一个问题就是云计算中心数据存储的安全。由于云计算的操作系统、数据处理技术以及存储的核心技术都在美国IT巨头手中，如果云端存在漏洞和后门，就无法确保在云端的数据安全。

（3）物联网技术

物联网基于下一代网络技术提供的通信服务和云计算中心提供的强大运算能力。物联网打破了之前网络通信的信息来源，目前其主要信息来源是人与物之间的交互信息。通过链路层的一些新创的通信协议，能够实现将地球上需要监控的实体信息联网，从而将原来虚拟的互联网空间变成了可感知的真实世界[16]。

基于美国IBM主导的“智慧地球”战略，其目的就是利用IBM掌握的核心技术资源，不断推出基于互联网的新技术、新产品和新应用。该战略的核心就是要建立网络社会，将现实世界的所有活动全部纳入互联网管理。随着现实生活对互联网的依赖，美国不仅可以掌握中国经济、政治、军事的信息，甚至可以随时实施信息干预和制裁。这不仅意味着中国的产业和经济安全无法得到保障，还意味着政府的执政能力也将受到来自美国IT巨头的严重挑战。

物联网的应用使得最为封闭的电力网络也进入了互联网时代。由于电力网络关乎百姓的日常生活和国民经济的生产活动，如果电力网络遭到攻击，会给国民经济造成重大的损失。

（4）智能手机即时通信软件的相关应用

智能手机作为移动互联网终端

的重要组成部分，可以作为下一代信息技术研究的重点。传统基于PC机之上的QQ、微博等通信软件和互联网应用，正在被智能手机上的即时通信软件和其他APP所替代。以朋友圈和公众微信号为代表的应用能够大面积传播各种未经证实的信息，容易引起社会的恐慌，还有可能引发群体性事件。

3.2 应对策略

下一代信息技术对网络空间安全影响的应对策略主要有以下几个方面。

（1）加快下一代网络通信基础设施的部署

下一代网络技术在通信协议方面主要基于IPv6技术，并对网络层协议进行重大的改变。如果不借助过度技术，IPv6和IPv4则无法在网络层进行兼容。目前只在高等院校和科研院所存在着一些IPv6的信息孤岛。中国应该对下一代网络骨干网的建设投入巨额资金，争取在“十三五”期间完成大中城市基于下一代网络骨干网的建设，让中国在下一代网络的建设中走在世界大国的前列。

通过下一代网络的建设，能够让中国拥有足够多的独立IP地址，以推进物联网的应用。另外，通过下一代网络的建设还可以扶持以中兴通讯为代表的众多高科技网络通信公司，自主研发出高性能的下一代高速路由器和交换器，抢占世界市场。

（2）自行研制云计算的核心技术

云计算技术在中国方兴未艾，虽然中国拥有众多的云计算中心，但是这些中心缺乏核心的云计算技术，包括计算技术和存储技术。中国应该整合现有的云计算平台，为研发云计算核心技术提供开发、测试与实验提供可靠强大的平台。

（3）制订基于物联网技术的“互联网+”的战略工程

在国家“十三五”规划中，主推“互联网+”战略工程。在中国经济转型过程中，物联网技术能够为“互联网+”工程提供强有力的支撑平台和技术保证。

（4）对智能手机的即时通信进行有效管控

对基于智能手机的即时通信软件推行实名制，对朋友圈和公众号的的发布信息进行实时监控，对敏感的关键词进行有效过滤与实时预警，防止不良信息的广泛传播。

（5）从国家战略的高度重视下一代信息技术标准的制订

制订下一代信息技术通信与应用标准，例如移动通信5G技术和移动支付标准等，能够为中国经济带来巨额的收入。例如，美国高通公司通过标准和专利，每年高达80亿美元的利润。

（6）大力支持国家下一代信息技术自主创新支撑体系

推进国家高等院校和科研院所协作建立下一代信息技术自主创新体系和标准，并建立下一代信息技术研发和测试平台，为中国下一代信息产业的发展和应用提供强有力的物质支撑。

以应用推进下一代技术向广度和深度二维方向发展，以应用推进技术的发展，推进更广泛的用户的应用，使得技术发展与用户的应用进入良性循环，这是所有信息技术强国的必由之路。

4 结束语

网络环境的复杂性、多变性，以及信息系统的脆弱性，决定了网络空间安全威胁的客观存在。随着中国的日益开放，网络空间安全监管的加强和保护屏障的建立变得不可或缺。网络空间安全是涉及中国经济发展、社会发展和国家安全的重大问题。通过自主可控和创新才能从技术上应对网络空间安全面临的威胁与挑战。中国应该从战略高度重视下一代网络技术对网络空间主权的威胁与挑战，并采用相应的技术与政策及时应对。

参考文献

[1] 丁禹, 逯志安, 焦建伟. 2014年网络空间安全问题综述与展望[J]. 通信安全与信息保密, 2015, (2): 16–21
[2] 田力加, 王光厚. 中国网络空间安全现状研究[J]. 山西大同大学学报, 2015, 29(2): 12–14
[3] 廖东升, 石海明, 郭勤, 等. 全球视阈下的网络空间国家安全战略[J]. 湖南社会科学, 2013, (6): 43
[4] 何德旭, 饶云清, 王智杰. 金融安全网: 基于信息空间理论的分析[J]. 经济理论与经济管理, 2011, (2): 69–78
[5] 王世伟. 论信息安全、网络安全和网络空间安全[J]. 中国图书馆学报, 2015, (2): 72–84
[6] 惠志斌.我国国家网络空间安全战略的理论构建与实现路径[J]. 中国软科学, 2012, (5): 22–27
[7] 彭长艳. 空间网络安全关键技术研究[D]. 长沙: 国防科学技术大学, 2010
[8] 林伟. 空间网络技术的研究与实现[D]. 成都: 电子科技大学, 2012
[9] 张钢. 网络空间安全问题探讨[J]. 科技信息, 2013, (5): 96–97
[10] 方兴东, 张笑容, 胡怀亮. 棱镜门事件与全球网络空间安全战略研究[J]. 现代传播: 中国传媒大学学报, 2014, 36(1): 115–122
[11] 雷璟. 网络空间攻防对抗技术及其系统实现方案[J]. 电讯技术, 2013, (11) : 1494–1499
[12] 董淑英. 探究网络空间的自主构建与管理[J]. 信息安全与通信保密, 2013, (9): 47–52
[13] 王鹤鸣. 网络安全新政一实体战争与网络战争的突袭[J]. 信息安全与通信保密, 2011,(7): 15–16
[14] 胡连宛.中国网络安全面临严峻挑战[J]. 决策与信息, 2012, 330(5): 4–8
[15] 曲成义. 网络空间安全保密对抗态势和应对策略[J]. 电讯技术, 2012, (4): 6–7
[16]童淑英. 物联网与网络空间安全[J]. 河北省科学院院报, 2011, (3): 77–81

作者简介

周延森，国际关系学院信息科技学院信息安全教研室主任、副教授；主要研究方向为网络通信及安全、智能手机应用及安全；先后主持完成10多个校级项目；发表论文10余篇。

周琳娜，国际关系学院信息科技学院常务副院长、电子与通信工程学科组和警务科技学科组组长，国家百千万工程人才，国家重点领域创新团队负责人；主要研究方向为信息安全、数字取证、多媒体信息处理等；先后申请和主持了国家自然科学基金等重点项目共4项，获得国家科技进步一、二等奖各1项，获国家技术发明奖1项，获得省部级科技进步奖30余项；出版专著3本，发表SCI、EI检索论文30余篇。

DOI:10.3969/j.issn.1009-6868.2016.01.003
网络出版地址:http://www.cnki.net/kcms/detail/34.1228.TN.20151204.1604.002.html

网络空间安全体系及关键技术

Security Architecture and Key Techniques of Cyberspace

张应辉/ ZHANG Yinghui[1]
郑东/ ZHENG Dong[1]
马春光/ MA Chunguang[2]
(1. 西安邮电大学 通信与信息工程学院,陕西 西安 710121;
2. 哈尔滨工程大学 计算机学院,黑龙江 哈尔滨 150001)
(1. School of Telecommunication and Information Engineering, Xi'an University of Posts and Telecommunications, Xi'an 710121, China;
2. College of Computer Science, Harbin Engineering University, Harbin 150001, China)

中图分类号:TN929.5 文献标志码:A 文章编号:1009-6868 (2016) 01-0010-004

摘要: 提出一种新型立体式网络空间安全体系结构,新结构有助于实现立体式网络空间安全防御体系,克服了传统线性防御体系只能应对单一性安全威胁的缺点。在新的体系结构中,网络空间中的节点分布于所有层次之中,且每一层的活动支撑着其他层的活动,并对整个网络空间产生影响。此外,针对所提出的立体式网络空间安全体系结构,结合当前的热点应用,指出了立体式网络空间安全防御体系应采取的关键技术。

关键词: 网络空间;立体式安全体系;信息安全

Abstract: A novel three-dimensional security architecture of cyberspace is proposed. This architecture has contributed to the construction of the three-dimensional security defensive system of cyberspace, and it overcomes the disadvantage of tackling only unitary security threats in the traditional linear defense system. In the proposed architecture, each node is involved in all layers in cyberspace, and the events in a layer affect other layers and thus affect the whole cyberspace. In addition, considering the proposed three-dimensional security architecture, security techniques suitable for the three-dimensional security defensive system are pointed out. These techniques are based on the current hot applications in practice.

Key words: cyberspace; three-dimensional security architecture; information security

网络空间[1]一词最早出现在美国科幻小说中,在故事中主角将自己意识感知的世界称为网络空间。随着时代的发展,网络空间被赋予了很多新的含义。网络空间是连接各种信息技术基础设施的网络,包括互联网、各种计算机系统及人与人之间相互影响的虚拟环境。

网络不仅是一个消息的载体和媒介,它还改变了我们周围的一切,并悄悄地改变着我们的思维[2]。从某种程度上讲,人们所处的环境,都被赋予了网络和信息的属性[3]。因此,我们可以认为网络空间安全[4]的核心是信息安全。如今,信息技术以及其工业应用迎来了前所未有的繁荣,信息安全问题也变得越来越突出。此外,科学与技术的发展给信息安全带来了新的挑战,利用量子[5]与DNA计算,许多现存的公钥加密系统变得不再安全,网络空间的安全问题变得越来越严峻。

收稿日期:2015-10-23
网络出版时间:2015-12-04

1 网络空间安全分析

在给出网络空间安全体系结构之前,我们首先对网络空间的安全性进行分析。

1.1 传统意义上的网络空间安全范畴

1.1.1 物理电子设备安全

物理电子设备是我们存放消息数据的载体,从这个意义上讲,对于其安全性的考虑,不仅包括硬件电子设备在硬件上的不被恶意损毁和盗取,也应包括用户存放于上面的数据不应被人为地通过物理手段窃取或者删除。如何保证存放数据的物理设备有一定的灾备能力,如何从毁坏的设备中恢复用户的数据等问题都非常重要。此外,旁路攻击[6]中利用电磁信号变化、电位变化等,通过统计学恢复加密数据明文的攻击手段也属于此范畴。

1.1.2 应用层与系统层安全

应用层与系统安全主要是指当用户的数据在计算机系统中储存的时候,系统和应用层是安全可靠的。这里的安全威胁主要来自于不可信系统,或者恶意应用软件。

1.1.3 网络安全

传统意义上的网络安全是如何保障网络互连互通安全。网络互连指的是将不同的网络通过连接设备连接形成一个巨大的网络,或者是为

了便于管理，将一个更大的网络划分为几个子网，而网络互通是指建立各个子系统共享资源的环境，网络互连比网络互通更易实现。为了实现网络互连，一般使用中继器、网桥、路由器等设备，而为了实现网络互通，必须考虑各个子系统之间数据交流协调同步问题，同时需要设置各个子系统之间硬件和软件的参数等。如何在网络空间上安全传播，不被恶意窃听和修改，便成了重中之重。网络传播过程的安全，也不仅仅指的是传播链路上的安全，还应包括提供网际传输链路的服务提供商在硬件和软件上，不被恶意攻击，链路可以畅通无阻。针对网络的常见攻击手段有拒绝服务攻击、中间人攻击等。

1.1.4 人员管理安全

网络管理人员是网络空间的一个至关重要的组成部分，管理人员依靠专业知识规划、监督、控制着网络资源的使用以及网络中的各种活动，从而使得网络的安全性能达到最优化。因此，从信息技术角度上解决网络安全问题的同时，我们必须加强对网络网络管理人员的监督以及管理。网络管理人员对网络安全的威胁不仅包括管理人员的监守自盗、擅权越权等非法操作，更包括安全意识薄弱、管理环节不健全等潜在威胁。所以我们要采取切实可行的措施，制订更加严格的管理制度，不断提高和加强网络管理人员的管理水平以及安全意识。

1.2 新形势下的网络空间安全范畴

近几年，随着智能移动终端的普及，人们的生活与网络联系更为密切。因此，除了4个传统意义上的安全考虑之外，新的形势下还有很多新的网络空间安全问题。这里所谓的新形势是指：在当代新技术不断涌现，各领域高度融合的前提下，网络空间安全所展现出来的新局面。新形势之所以新，是因为：

（1）所处层次的复杂化。如果不考虑人员层的管理安全，以往的安全问题出现的时候，所处层次往往比较单一，比如上述提到的XcodeGhost事件，就出现在编译环境和由该环境生成的代码中，属于应用层范畴。但是新形势下的网络空间安全往往是跨区域、跨层次的。

（2）表现形式的多元化。传统安全问题表现形式较为单一，比如个人隐私，在传统的思维模式里，用户可以根据自己意愿对所持有数据进行公开。但是在大数据时代，数据挖掘技术和机器学习学习技术能从用户已公开的数据中嗅探出用户不愿意公开的数据，隐私的表现形式已经不仅仅是自己不愿意公开的数据，更广泛地分布在已经公开的碎片数据中。

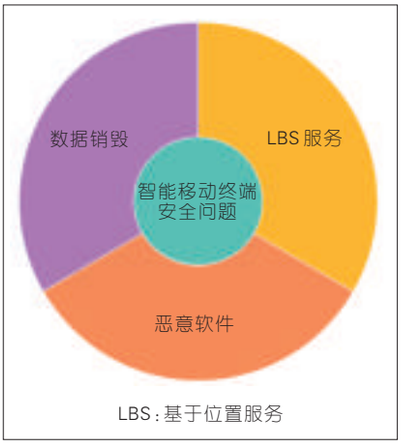
（3）涉及技术的多样化、新型化。比如下面提到的物联网技术，它涵盖了材料技术、生物技术、通信技术等，每个环节都有新的安全指标和挑战。很多技术在以往的安全关键技术中都是很少涉及的，比如材料技术。涉及到安全问题的一个典型例子就是最近逐渐兴起的可编程材料技术，工程师可根据需求对材料进行编程，改变材料结构。那么，我们可以断言，距离材料型病毒诞生的一天也不远了。

可以看出，新形势下的安全问题是对传统问题的延续与补充，传统问题所表达的安全基准在新形势下也同样适用。新形势下的安全问题很可能是若干传统安全问题的交集。下面我们就结合当下的热点应用，阐述新形势下的安全问题。

1.2.1 智能移动终端安全

近年来，手机等智能移动终端迅速发展，很多安全问题也随之暴漏出来，如图1所示，具体包括：

（1）恶意软件。智能终端的恶意软件和PC端的恶意软件具有同样的危害。从所属层次上来说，终端恶意软件仍然属于应用层和系统层层面，



▲图1 智能移动终端安全

但是由于移动终端的存储能力和计算能力有限，终端恶意软件多以后门、木马的形态存在。所以，终端恶意应用正逐渐向网络层过渡。

（2）基于位置服务[7]。基于位置服务是指通过运营商或者外部设备获取移动终端设备位置信息的服务。如何保证一个基于位置服务提供商是可信的，不会将用户的位置信息暴漏给其他第三方，是值得考虑的一个问题。基于位置的服务是网络层、应用层与物理设备层的相互交叉的产物。

（3）数据销毁。当更换手机或硬盘时，人们会把旧的设备格式化，以清除数据，避免信息泄露。数据销毁则是需要物理设备层、应用层与系统层协调工作。

1.2.2 可穿戴设备安全

近几年还有一些其他类型嵌入式系统和可穿戴设备的安全性也引起了人们的重视。常见的可穿戴设备是指那些具有部分计算能力，与智能移动设备相辅使用的便携式设备。这些设备多以手表、鞋子、帽子等形式存在，边缘化的还有一些服装、书包、配饰等。然而，在2015年的HackPWN安全极客狂欢节上，有白帽子黑客向组委会递交了一个小米手环的漏洞，通过该漏洞，黑客可以完全接管小米手环的控制权。要

想解决可穿戴设备安全问题，应该从物理设备层与系统层进行考虑。

1.2.3 云计算安全

云计算[8]在近几年受到了学术界、产业界和政府等的共同关注。云计算的安全主要包括：

（1）虚拟化安全。虚拟化技术在信息系统中发挥着极其重要的作用，它可以降低信息系统的操作代价、改进硬件资源的利用率和灵活性。但随着虚拟技术的广泛运用，其安全问题越来越受到人们的关注。

（2）云存储安全。云存储可以为用户提供海量的存储能力，而且可以减少成本投入。然而，由于对数据安全性的担忧，仍然有很多用户不愿意使用云存储服务。如何保证用户所存储数据的私密性，完整性等都是云存储安全的范畴。

1.2.4 物联网安全

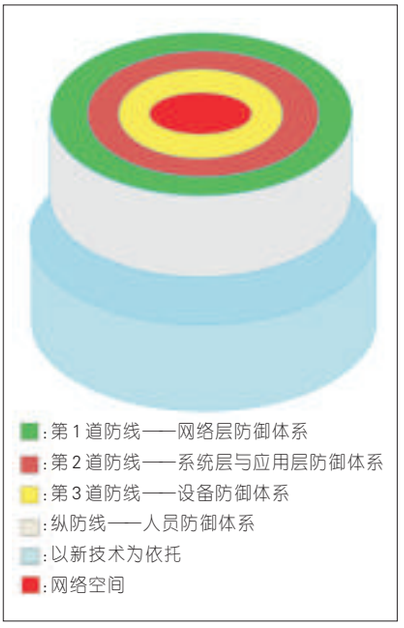
物联网[9]被视为继计算机、互联网和移动通信之后的第3次技术革命和信息产业浪潮，它广阔的行业前景和潜在的巨大市场规模受到了各国政府和研究者的极度重视。物联网涵盖了材料技术、生物技术、计算机技术、电子技术、通信技术，打破了行业之间的界限，实现了通信从人与人向人与物，甚至于物与物之间拓展。然而，也正因为如此，物联网的安全才更加具有挑战性。

1.2.5 量子计算机对传统密码学算法带来的挑战

随着科学的进步与发展，诞生了很多新兴的技术，如量子计算机技术。量子计算机的诞生，可能对传统意义上的密码学构成威胁，其特点是计算能力非比寻常，将在现有计算能力上实现指数增长。目前来说，量子计算机还处于萌芽期，不具备可操作性，而且实验性量子计算机也不足以对传统加密算法发起攻击，但是随着政府资金的大量投入，理论和实践活动的开展，实用性量子计算机或许随时都会诞生。传统密码算法[10-11]所依赖的大整数分解，椭圆曲线以及离散对数问题在大规模量子计算机面前，会变得不堪一击。

2 网络空间安全体系结构

随着中国在电子银行、电子商务和电子政务方面的进一步发展，网络空间安全形势也越来越严峻，急需构建新型网络空间安全防御体系，从传统线性防御体系向新型立体式网络空间防御体系发展。传统线性防御体系只能应对某一层内的单一性安全威胁，而新型网络空间防御体系必须能够实现立体化的安全防域，即该网络空间中的节点分布于所有域之中，网络空间中的所有活动支撑着其他域中的活动，并且其他域中的活动同样能够对网络空间产生影响。立体式网络空间安全体系如图2所示。



▲图2 立体式网络空间安全防御体系

2.1 第1道防线——网络层防御体系

网络防御层[12]是保证信息数据在网络传输过程中的坚固堡垒与屏障。防御模式不仅包括传统意义上的虚拟专用网络（VPN）、防火墙等。而且为了确保网络互连互通安全，需要加强对网络互连互通设备的安全设置，比如中继器、网桥、路由器等等。在新兴技术的支持下，网络层安全如虎添翼。服务器可以通过固化在用户终端的安全模块，对用户的网页浏览行为进行管控，确保用户所浏览的网页没有受到钓鱼网站的劫持冒用。同时，用户之间的会话在逻辑上是加密的，并且会话密钥的分享方式是安全可靠的。

2.2 第2道防线——系统层与应用层防御体系

系统层与应用层防御是针对于软件而言的，在系统最先开始编译的时候，就可以将一些安防软件、病毒检测软件内嵌到系统中，尤其是在软件开源的大趋势下，我们完全有能力将安防体系作为系统模块的一部分，固化在操作系统本身的结构中。面向的对象，包括一些底层是Linux嵌入式系统的可穿戴设备，以及拥有开源优势的安卓手机系统，都可以被这个防御思维有针对性地进行改造。在用户接入系统的时候，通过系统将对所有的用户进行约束与管控：通过在终端上部署防病毒客户端，有效控制病毒的感染与传播，依托于大数据云计算平台，进行终端和网络病毒查杀，以保证计算终端配置和软硬件信息不被恶意病毒修改。通过主机对终端硬件如磁盘、外接设备等的安全监控，可以实现移动存储介质的安全接入，控制终端用户对核心系统的读写等。

2.3 第3道防线——设备层防御体系

设备层防御体系是从硬件上构建的防御体系，从最底层打牢网络空间的基石。通过改良硬件的基础设施，在硬件最初被设计的时候，就将其安全功能考虑进去，必要时候可以在芯片中内嵌一些安全算法，布控一些安全防御设备，包括反窃听、反旁路攻击等。合理规划安排硬件安装

过程中的每一个环节，对硬件的操作进行软件上或者物理上的监控。确保在设备层面构成网络空间的第3道防线。

2.4 纵防线——人员防御体系

建设人员防御体系的核心是建设合理的人员管理体系。对人员的安防意识进行有针对性的强化，同时加强道德品质建设，对于具有专业能力的计算机从业人员进行正确疏导，避免误入歧途。加强法律的威慑力与约束能力，加强安防软件硬件的基础设施建设，加快落实实名制，在实名制的基础上引入生物特征识别机制，加大网络犯罪的犯罪成本。人员的管理穿插在防御体系的每一个环节中，因此非常值得关注。

3 网络空间安全关键技术

3.1 智能移动终端恶意代码检测技术

针对智能移动终端恶意代码而研制的新型恶意代码检测技术，是在原有PC机已有的恶意代码检测技术的基础上，结合智能移动终端自身的特点而引入的新技术。从检测方法上，可分为动态监测和静态检测。因为智能移动终端自身的计算能力有限，手机端恶意代码检测往往需要云查杀辅助进行。与手机数据销毁相对应，手机取证[13]也有着极为重要的应用。手机取证是打击犯罪的重要手段之一。在手机取证中，手机的SIM卡、内外存设备，以及手机所对应的服务提供商都是手机取证的重要环节。

3.2 可穿戴设备安全防护技术

3.2.1 生物特征识别技术

英特尔首席执行官的科在2014年的CES预热演讲中强调，英特尔将推出“Intel Security”品牌，用安全领军可穿戴设备。讲话中还提到，英特尔将把生物特征识别技术应用于可穿戴设备中。生物特征识别技术指的是用生物体本身的特征对一个人进行身份验证，这其中的一些技术，比如指纹识别技术，已经被用户们所熟知。除此之外，近年来又新兴了如步态识别、脸像识别、多模态识别技术的新一代生物特征识别技术。可穿戴设备可以对用户的身份进行验证，如果验证不通过将不予提供服务。

3.2.2 入侵检测与病毒防御工具

保证可穿戴设备安全的另一个重要思路就是在设备中引入入侵检测与病毒防护模块。由于可穿戴设备中本身的计算能力非常有限，所以，嵌入在可穿戴设备中的入侵检测或者病毒防护模块只能以数据收集为主，可穿戴设备通过网络或者蓝牙将自身关键节点的数据传递到主控终端上，再由主控终端分析出结果，或者通过主控终端进一步转递到云平台，最终反馈给可穿戴设备，实现对入侵行为或者病毒感染行为的发觉与制止。

3.3 云存储安全技术

3.3.1 云容灾技术

利用物理上隔离的两台设备以及一些特殊的算法，实现资源的异地分配。当有一台或者数台物理设备被意外损毁，用户仍然可以通过储存在其他设备上的冗余信息恢复出原数据。比较有代表性的就是基于Hadoop的云存储平台，其核心技术是分布式文件系统（HDFS）。在硬件上，云容灾技术不依赖具体的某一台物理设备，并且不受地理位置的限制，使用非常方便。未来应进一步考虑效率更高、更稳定的云容灾技术。

3.3.2 可搜索加密与数据完整性校验技术

用户可以通过关键字搜索云端的密文数据。新的可搜索加密技术应该关注关键词的保护，支持模糊搜索，即允许用户在搜索的时候输入错误，同时还要支持多关键词检索，并对服务器返回的结果进行有效性验证。此外，为了实现数据的完整性验证，使得用户并不需要去完全下载自己存储在云端的数据，而是基于服务器提供的证明数据和自己本地的小部分后台数据。未来新的完整性审计技术应该支持用户对数据的更新，同时保证数据的机密性。

3.3.3 基于属性的加密技术

基于属性的加密[14]是一种非常有吸引力的密码学原语，支持一对多加密模式。在基于属性的加密系统中，用户向属性中心提供属性列表信息或者访问结构，属性中心返回给用户私钥。数据拥有者选择属性列表或者访问结构对数据进行加密，把密文外包给云服务器进行存储。在基于属性的环境中，由于不同的用户可以拥有相同的属性信息，因此可以具有一样的解密能力，从而导致属性撤销和密钥滥用的追踪问题。未来基于属性的加密技术应同时考虑密钥滥用的追踪和属性撤销机制。

3.4 后量子密码

现代密码学是建立在计算复杂性理论基础之上的。然而，量子计算机的高度并行计算能力，可以将相应的困难问题化解为可求解问题。以量子计算复杂度为基础设计的密码系统必然具有抗量子计算的性质，从而有效地增强了现代密码体制的安全防护。未来量子密码的研究主要还应关注实用的量子秘钥分发协议。此外，编码密码技术也具有抵抗量子算法攻击的优点，是信息技术领域不可缺少的重要技术之一。未来的研究可以关注基于编码的加密技术、基于编码的数字签名技术等。

4 结束语

随着计算机网络技术的快速发

➡下转第18页

DOI:10.3969/j.issn.1009-6868.2016.01.004
网络出版地址:http://www.cnki.net/kcms/detail/34.1228.TN.20151117.1514.012.html

云时代下的大数据安全技术

Security Technology of Big Data in the Cloud Era

杨曦/YANG Xi [1,2]
GUL Jabeen [1]
罗平/LUO Ping [1]

(1. 清华大学 信息系统安全教育部重点实验室,北京 100084;
2. 福州大学 阳光学院计算机工程系,福州 350015)
(1.The Key Laboratory for Information System Security, Ministry of Education, Tsinghua University, Beijing, 100084, China;
2. Computer Engineering Department, Sushine College of Fuzhou University, Fuzhou 350015, China)

中图分类号:TN929.5 文献标志码:A 文章编号:1009-6868 (2016) 01-0014-005

摘要: 认为云计算结合大数据,是时代发展的必然趋势。提出了保障大数据安全的方法和技术,方法包括:构建云环境下的大数据信息安全体系,建立并研究基于Hadoop的大数据安全架构等;技术包括:基于大数据的威胁发现技术、大数据真实性分析技术、基于大数据的认证技术、基于大数据的安全规则挖掘技术,以及防范高级持续性威胁(APT)攻击的技术。认为大数据带来许多新的安全问题和挑战,但它本身也是解决问题的重要手段,需要进一步地研究。

关键词: 大数据;云计算;大数据安全;APT攻击;数据挖掘

Abstract: The combination of cloud computing and big data is an inevitable trend. In this paper, methods and techniques for ensuring the security of large data are presented. These methods include: building a large data information security system in a cloud environment and establishing and studying the big data security architecture based on Hadoop. These techniques include: threat discovery based on big data, big data authenticity analysis, authentication based on big data, security rule mining based on big data, and preventing advanced persistent threat (APT) attack. Big data creates many new security problems and challenges, but it is also an important means to solving the problem, which needs for further research.

Key words: big data; cloud computing; big data security; APT attack; data mining technology

随着云时代的来临,大数据也吸引了越来越多学术界和工业界的关注。从20世纪90年代“数据仓库之父”Bill Inmon率先提出“大数据”的概念,到2011年麦肯锡全球研究院(MGI)发布了关于大数据的详尽报告,直至2012年美国奥巴马政府公布了“大数据研发计划”,才使得大数据真正成为许多学科的重点研究课题。大数据科学的基础研究已经成为当今社会的研究热点。英国牛津大学教授维克托·迈尔·舍恩伯格,在他的《大数据时代:生活、工作与思维的大变革》一书中,深刻地阐述了大数据所带来的三大变革,即思维变革、商业变革和管理变革。大数据带来更多的是思维变革——样本数据或局部数据向全体数据的变革,结果数据向过程数据的变革,静态存储数据向动态流处理数据的变革。

收稿日期:2015-11-10
网络出版时间:2015-11-17
基金项目:国家自然科学基金(60973142);福建省教育厅A类科技项目(JA14358)

随着大数据技术的不断发展,许多传统的信息安全技术也受到了挑战。在大量数据产生、收集、存储和分析的过程中,既会涉及一些传统安全问题,也会涉及一些新的安全问题,并且这两类问题会随着数据规模、处理过程、安全要求等因素而被不断放大。而大数据的4V(大量、高速、多样、真实性)+1C(复杂)特征,也使得大数据在安全技术、管理等方面面临新的安全威胁与挑战[1]。

1 大数据安全技术发展现状

谈到大数据,不可避免地就要提及云计算技术,它们就像一枚硬币的正反面一样密不可分。云计算结合大数据,是时代发展的必然趋势。云计算为大数据提供了存储场所、访问渠道、虚拟化的数据处理空间,具有盘活数据资产价值的能力。另一方面,大数据技术通过挖掘价值信息[2]进行预测分析、策略决断,为国家、企业甚至个人提供决策和服务。

作为一个云化的大数据架构平台,Hadoop自身也存在着云计算面临的安全风险,企业需要实施基于身份验证的安全访问机制,而Hadoop派生的新数据集也同样面临着数据加密问题。云端大数据从使用频率上有静态数据加密机制和动态数据加密机制两种[3]。静态数据加密机制与传统加密一样,有对称加密算法和非对

称加密算法两种。而动态数据加密机制方面近年来则有较多的论述，较为常用的是同态加密机制[4]。对加法同态的加密算法有 Paillier 算法[5]，对乘法同态的加密算法有 RSA 算法，还有对加法和简单标量乘法同态的加密算法，如 IHC 和 MRS 算法[6]。Craig Gentry 提出一种基于理想格的全同态加密算法[7]，实现了全同态加密所有属性的解决方案。

同样，大数据依托的非关系型数据库（NoSQL）技术没有经过长期发展和完善，在维护数据安全方面也未设置严格的访问控制和隐私管理，缺乏保密性和完整性特质。另一方面，NoSQL 对来自不同系统、不同应用程序及不同活动的数据进行关联，也加大了隐私泄露的风险。大数据时代，想屏蔽外部数据商挖掘个人信息是不可能的，大数据隐私问题堪忧。Itani 提出的协议能够在云计算环境下保证用户的隐私[8]，Creese 的方案有效地解决了企业云部署中的隐私安全问题[9]。除了常见的基于加密体制的数据存储和数据处理的隐私性保护方案外，A. Parakh 等于 2011 年和 2013 年分别提出了基于空间有效性的机密共享隐式机制[10]及运用隐式机制的云端计算机制[11]。针对非结构化数据（比如社交网络产生的大量数据）的隐私保护技术也是云时代下大数据安全隐私保护的重大挑战，典型的匿名保护需求为用户标识匿名、属性匿名（也称点匿名）及边匿名（用户间关系匿名）。目前边匿名方案大多是基于边的增删[12]，还有一个重要思路是基于超级节点对图结构进行分割和聚集操作[13]。

2 基于大数据的安全技术及发展趋势

新形势下的大数据安全也面临诸多新的挑战，在大数据产业链的各个环节，安全问题无处不在。面对一系列的安全风险和关键问题，如何保障大数据安全，并在信息安全领域有效利用，是学术界和工业界都需要认真对待和解决的问题。

2.1 构建云环境下的大数据信息安全体系

只有在正确完整的安全体系指导下，大数据信息安全建设所需的技术、产品、人员和操作等才能真正发挥各自的效力。大数据应用过程通常划分为采集、存储、挖掘、发布 4 个环节，它们的安全性可通过下面一些技术和方法实现：

（1）数据采集阶段的安全问题主要是数据汇聚过程中的传输安全问题，需要使用身份认证、数据加密、完整性保护等安全机制来保证采集过程的安全性。传输安全主要用到虚拟专用网络（VPN）和基于安全套接层协议 VPN（SSL VPN）技术。

（2）数据存储阶段需要保证数据的机密性和可用性，提供隐私保护、备份与恢复技术等。这个阶段可能用到的技术有：基于数据变换的隐私保护技术（包括随机化、数据交换、添加噪声等）、基于数据加密的隐私保护技术、基于匿名化的隐私保护技术（通常采用抑制、泛化两种基本操作）、静态数据加密机制（数据加密标准（DES）、高级加密标准（AES）、IDEA、RSA、ElGamal 等）、动态数据加密机制（同态加密）、异地备份、磁盘阵列（RAID）、数据镜像、Hadoop 分布式文件系统（HDFS）等。

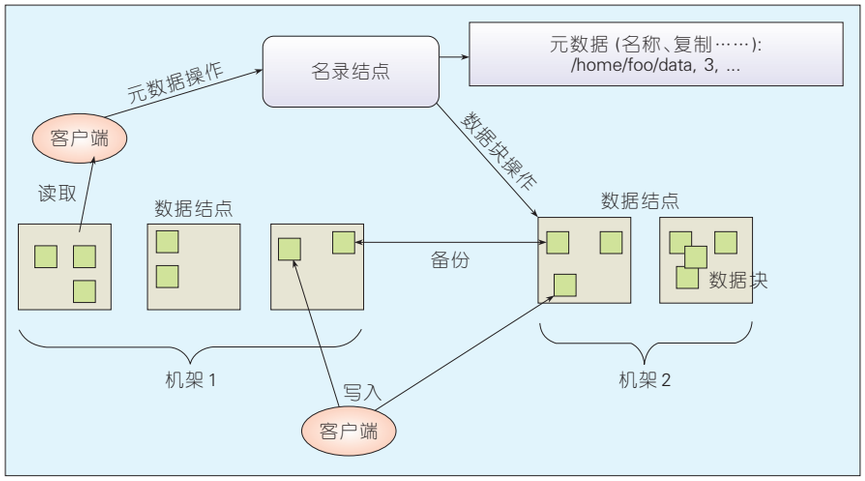
（3）数据挖掘阶段需要认证挖掘者的身份、严格控制挖掘的操作权限，防止机密信息的泄露。这个阶段涉及到的技术有：基于秘密信息的身份认证、基于信物的身份认证技术、基于生物特征的身份认证技术、自主访问控制、强制访问控制、基于角色的访问控制等。

（4）数据发布阶段需要进行安全审计，并保证可以对可能的机密泄露进行数据溯源。这个阶段的技术可能涉及到：基于日志的审计技术、基于网络监听的审计技术、基于网关的审计技术、基于代理的审计技术、数据水印技术等。

2.2 基于 Hadoop 的大数据安全架构

Hadoop 是一种分布式数据和计算的框架，在全球范围内已成为大数据应用最为广泛的技术架构。当前，Hadoop 已成为工业界和学术界进行云计算应用和研究的标准平台。分布式文件系统使大规模并行计算成为可能，但堆栈各层的集成以及数据节点与客户端/资源管理机构之间通信，都会引入新的安全问题。图 1 是 Hadoop 核心 HDFS 的架构，在不破坏大数据集群的基本功能及大数据本身必要特点的前提下，我们先来分析这种架构下的安全问题及隐患并给出相应安全解决建议。

在高度分布式数据集群中，很难验证异构平台之间安全的一致性，即不同的数据结点的数据安全的整体性和一致性是分布式计算的痛点。而与传统集中式数据安全模型不同的是，大数据集群内的数据是流动的，有多个副本，在不同节点间移动以确保冗余和弹性的机制导致数据很难及时、准确地定位存储位置，无法获知数据备份个数，这加大了副本安全保护机制设计上的难度。对于数据访问，大多数大数据环境提供了 schema 级别的访问控制，但没有更细的粒度，虽然在大数据环境中可以借鉴安全标签和其他高级属性，但需要应用设计者将这些功能集成到应用和数据存储中去。对于节点间的通信，Hadoop 和绝大多数组件之间的通信是不安全的，它们使用传输控制协议（TCP）/IP 之上的远程过程调用协议（RPC），并没有嵌入安全传输层协议（TSL）和 SSL 等安全机制。另外，客户端可以直接与资源管理者及节点进行交互，增加了恶意代码或链接发送的概率，也难以保证客户端免受数据节点的攻击。最后，最为重要的是大数据栈自身设计并没有考虑安全机制。这些都是基于 HDFS 架构的



▲图 1 Hadoop 的 HDFS 架构

大数据环境的安全隐忧。

基于 Hadoop 的大数据架构，其安全机制可以通过下面一些方法和技术得以保证：

（1）使用 Kerberos 进行节点验证。Kerberos 是一个最有效的安全控制措施之一，并且可以集成到 Hadoop 基础设施中。其可有效验证服务间通信，阻断集群中的恶意节点和应用程序，保护 Web 控制台的访问，使得管理通道难以被攻击。

（2）对于恶意客户端发起的获取文件请求，可以通过使用文件层加密对数据加以保护。被恶意访问的文件是不可读的磁盘映像，且文件层加密提供一致安全保护，有些产品甚至提供内存加密保护。

（3）使用密钥管理服务分发密钥和证书，并为每个组应用程序和用户设置不同密钥，可以提高密钥的安全性，防止文件加密的失效。

（4）在节点之间、节点与应用程序之间使用 SSL/TLS 组件实现安全通信，设计、集成有效的安全通信机制和现成组件。

2.3 基于大数据的威胁发现技术

由于大数据分析技术的出现，企业可以超越以往的“保护—检测—响应—恢复”（PDDR）模式，更主动地发现潜在的安全威胁。“棱镜”计划也可以被理解为应用大数据方法进行安全分析的成功故事。通过收集各个国家各种类型的数据，利用安全威胁数据和安全分析形成系统方法发现潜在危险局势，在攻击发生之前识别威胁。基于大数据的威胁发现技术可以使分析内容的范围更大，通过在威胁检测方面引入大数据分析技术，可以更全面地发现针对企业数据资产、软件资产、实物资产、人员资产、服务资产和其他为业务提供支持的无形资产等各种信息资产的攻击。另一方面，基于大数据的威胁发现技术可以使分析内容的时间跨度更长，现有的威胁分析技术通常受限于内存大小，无法应对持续性和潜伏性攻击。而引入大数据分析技术后，威胁分析窗口可以横跨若干年的数据，因此威胁发现能力更强，可以有效应对高级持续性威胁（APT）类攻击。基于大数据的威胁分析，我们可以对攻击威胁进行超前预判，能够寻找潜在的安全威胁，对未发生的攻击行为进行预防。而传统的安全防护技术或工具大多是在攻击发生后对攻击行为进行分析和归类，并做出响应。传统的威胁分析通常是由经验丰富的专业人员根据企业需求和实际情况展开，然而这种威胁分析的结果很大程度上依赖于个人经验。同时，分析所发现的威胁也是已知的。大数据分析的特点是侧重于普通的关联分析，而不侧重因果分析，因此通过采用恰当的分析模型可发现未知威胁。

2.4 大数据真实性分析技术

目前，基于大数据的数据真实性分析被广泛认为是最为有效的方法。基于大数据的数据真实性分析技术能够提高垃圾信息的鉴别能力。一方面，引入大数据分析可以获得更高的识别准确率。例如，对于点评网站的虚假评论，可以通过收集评论者的大量位置信息、评论内容、评论时间等进行分析，鉴别其评论的可靠性。如果某评论者为某品牌多个同类产品都发表了恶意评论，则其评论的真实性就值得怀疑。另一方面，在进行大数据分析时，通过机器学习技术可以发现更多具有新特征的垃圾信息。然而该技术仍然面临一些困难，主要是虚假信息的定义、分析模型的构建等。

云时代的未来必将涌现出更多、更丰富的安全应用和安全服务。对于绝大多数信息安全企业来说，更为现实的方式是通过某种方式获得大数据服务，结合自己的技术特色领域，对外提供安全服务。一种未来的发展前景是：以底层大数据服务为基础，各个企业之间组成相互依赖、相互支撑的信息安全服务体系，总体上可以形成信息安全产业界的良好生态环境。

2.5 基于大数据的认证技术

传统的认证技术主要通过用户所知的秘密（例如口令），或者持有的凭证（例如数字证书）来鉴别。这样就会存在问题：首先，攻击者总是能够找到方法来骗取用户所知的秘密或窃取用户持有的凭证，从而轻松通过认证；其次，传统认证技术中认证方式越安全往往意味着用户负担越重（例如携带硬件 USBKey），如果采用先进的生物认证技术，又需要设备具有生物特征识别功能，从而限制了

这些先进技术的使用。如果在认证技术中引入大数据分析则能够有效地解决这两个问题。基于大数据的认证技术指的是收集用户行为和设备行为数据，并对这些数据进行分析，获得用户行为和设备行为的特征，进而通过鉴别操作者行为及其设备行为来确定其身份。这与传统认证技术利用用户所知秘密、所持有凭证或具有的生物特征来确认其身份有很大不同。这样，攻击者很难模拟用户行为特征来通过认证，因此更加安全，同时又减小了用户认证负担，可以更好地支持各系统认证机制的统一。

2.6 基于大数据的安全规则挖掘技术

在Internet网络中，为保证网络安全，会引入防火墙技术和入侵检测技术等。在这些技术中，通常是通过建立一套安全规则或过滤规则达到其安全目标，而这些规则的建立传统方法是通过专家知识系统。在大数据时代，这些安全规则可以通过数据挖掘技术或方法实现。

聚类分析是数据挖掘中的一项重要技术，根据在数据中发现的描述对象及其关系的信息，将数据对象分组。组内相似性越大，组间差别越大，聚类效果就越好。

K-means算法作为聚类分析中的一种基本方法，由J.MacQueen于1967年首次提出[14]，由于其容易实现，时间复杂度与数据规模接近线性，并且能够快速地收敛到局部最优值，因此成为最广泛应用的聚类算法[15]。然而K-means算法也存在较为明显的缺陷，其中有以下两点：

（1）K-means算法需要人为确定聚类数K和选取初始质心集，其聚类结果的好坏明显受到初始化条件的影响[16-18]，即选取不同的K值和初始质心集会得到不同的聚类结果。

（2）K-means算法仅适用于数据项全是数字的情况。对非数字数据进行聚类分析是一个特别棘手的问题[19]，这在很大程度上限制了K-means算法的应用范围。

针对问题（1），Ester M等提出了基于密度的聚类方法DBSCAN[20]，该算法以及以此为基础的一些改进算法[17-18]采用基于密度的自动聚类，避免了对初始条件的随机选取，在一定程度上解决了K-means算法对初始条件敏感的问题。然而，由于基于密度的聚类算法时间复杂度通常较高，在处理大规模数据集时会出现瓶颈；同时在对于非数字数据集的聚类过程中，采用传统的基于密度的聚类算法往往会造成聚类失效问题。

针对以上问题，在借鉴K-means算法框架的基础上，文献[21]提出一种基于“预抽样-次质心”的密度聚类算法，采用预抽样的方法将算法时间复杂度控制为线性，同时通过引入次质心的概念，解决聚类失效问题。分析表明该算法能很好地克服K-means算法的初始条件敏感性和一般密度聚类算法的聚类失效问题，实现较为理想的聚类结果。

2.7 防范APT攻击的技术

APT攻击是大数据时代面临的最复杂的信息安全问题之一，而大数据分析技术又为对抗APT攻击提供了新的解决手段。APT具有极强的隐蔽性，且潜伏期长、持续性和目标性强，技术高级，威胁性也大。APT攻击检测方案通常有沙箱方案、异常检测、全流量审计、基于深层协议解析的异常识别、攻击溯源等。在APT攻击检测中，存在的问题包括：攻击过程包含路径和时序；攻击过程的大部分貌似正常操作；不是所有的异常操作都能立即检测；不能保证被检测到的异常在APT过程的开始或早期。基于早期记忆的检测可以有效缓解上述问题，既然APT是在很长时间发生的，我们的对抗也要在一个时间窗内来进行，并对长时间、全流量数据进行深度分析。APT攻击防范策略包括防范社会工程、通过全面采集行为记录避免内部监控盲点、IT系统异常行为检测等。

3 结束语

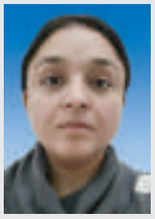
大数据带来许多新的安全问题和挑战，但大数据本身也是解决问题的重要手段，它就像一把双刃剑，既需要研究合适的“盾”来保护大数据，也需要研究如何用好大数据这根“矛”。战略资询公司麦肯锡认为：大数据将会是带动未来生产力发展、科技创新及消费需求增长的指向标，它以前所未有的速度，颠覆人们探索世界的方法，驱动产业间的融合与分立。大数据已成为各个国家和领域关注的重要战略资源，可能对国家治理模式、企业决策、组织业务流程、个人生活方式都将产生一系列长远、巨大的影响。

参考文献
[1] MANADHATA P K. Big Data for Security: Challenges, Opportunities, and Examples [C]// Proceedings of the 2012 ACM Workshop on Building Analysis Datasets and Gathering Experience Returns for Security, Raleigh, North Carolina, USA, 2012
[2] YU S C, WANG C, REN K, et al. Achieving Secure, Scalable, and Fine-grained Data Access Control in Cloud Computing [C]// Proceedings of the INFOCOM'10,the 29th conference on Information communication, Piscataway, USA, 2010: 534-542
[3] BELARE M and ROGAWAY P. Introduction to Modern Cryptography [J]. Ucsd Cse, 2005: 207
[4] GENTRY C, HALEVI S, SMART N P. Homomorphic Evaluation of The AES Circuit [M]. Germany: Springer Berlin Heidelberg, 2012
[5] CATALANO D. Paillier's Cryptosystem Revisited [C]// in Proceedings of the 8th ACM conference on Computer and Communications Security, PA, USA, 2001: 206-214
[6] BENDLIN R, DAMGARD I, ORLANDI C, et al. Semi-Homomorphic Encryption and Multiparty Computation [M]. Germany: Springer Berlin Heidelberg, 2011
[7] GENTRY C. A Fully Homomorphic Encryption Scheme [D]. Standford University, 2009
[8] ITANI W, KAYSSI A, CHEHAB A. Privacy As a Service: Privacy-Aware Data Storage and

Processing in Cloud Computing Architectures [C]// Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing, Washington DC, USA, 2009:711–716.doi: 10.1109/DASC.2009.139
[9] CREESE S, HOPKINS P, PEARSON S, et al. Data Protection–Aware Design for Cloud Services [M]. Germany: Springer Berlin Heidelberg, 2009
[10] PARAKH A, KAK S. Space Efficient Secret Sharing for Implicit Data Security [J]. Information Science, 2011, 181(2): 335–341
[11] PARAKH A, MAHONEY W. Privacy Preserving Computations Using Implicit Security [C] // Proceedings of the 22nd International Conference on Computer Communications and Networks (ICCCN), Nassau, Bahamas, 2013: 1–6. doi: 10.1109/ICCCN.2013.6614172
[12] ZHANG L J, ZHANG W N. Edge Anonymity in Social Network Graphs [C] // Proceedings of the International Conference on Computational Science and Engineering (CSE'09),Vancouver, Canada, 2009:1–8
[13] MICHAEL H A, GEROME M, DAVID J, et al. Resisting Structural Re–identification in Anonymized Social Networks[C] // Proceedings of the 34th International Conference on Very Large Data Bases (VLDB'2008), Aucklnad, New Zealand, 2008: 102–114
[14] MACQUEEN J. Some Methods for Classification and Analysis of Multivariate Observations. [C] // Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Oakland, USA, 1967: 281–297
[15] JAIN A K. Data Clustering: 50 Years Beyond K–Means [J]. Pattern recognition letters, 2010,1(8):651–666
[16] PIETRASZEK T, TANNER A. An Efficient K–Means with Good Initial Starting Points [J]. Georgian Electronic Scientific Journal: Computer Science and Telecommunications, 2009, 19(2):47–57
[17] SHEHROZ A A, KHAN S. Cluster Center Initialization Algorithm for K –Means Clustering[J]. Pattern Recognition Letters, 2004, 25(11): 1293–1302.doi: 10.1016/j.patrec.2004.04.007
[18] Stephen C H, REDMOND J. A Method for Initialising the K–Means Clustering Algorithm Using KD–Trees [J]. Pattern Recognition Letters, 2007, 28(8):965–973. doi: 10.1016/j.patrec.2007.01.001
[19] TAN P N, STEINBACH M, KUMAR V, et al. Introduction to Data Mining[J]. Pearson Addison Wesley Boston, 2006,1(1): 226–230
[20] ESTER M, KRIEGEL H P, SANDER J. A Density–Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise [J]. Kdd,1996, 96: 226–231
[21] GENGeng J K, YE DAREN, LUO P. A Novel Algorithm DBCAPSIC for Clustering Non–Numeric Data[C] // To Appear the ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Sydney, Australia, 2015

作者简介

杨曦，清华大学博士生；主要研究领域为软件可信性、软件工程理论与系统、数据库理论；先后主持和参加国家级基金项目5项，省部级项目10项；获得2项国家专利，发表论文10余篇。



GUL Jabeen，清华大学巴基斯坦籍博士生；主要研究领域为信息安全；已发表论文6篇，其中EI/SCI收录3篇。

罗平，清华大学教授；主要研究领域为网络空间安全；先后主持和参加国家级项目和国家科学基金项目20余项，获得教育部提名国家科学技术自然科学奖2等奖；已发表论文50余篇，其中被SCI检索30余篇。

⬅上接第13页

展，网络空间安全正面临着前所未有的发展机遇与挑战。通过分析传统线性结构防御体系以及传统的网络空间安全问题，结合当今严峻的新型网络空间安全威胁，我们提出一种新型立体式网络空间安全体系结构。新结构有助于实现立体式网络空间安全防域体系，并克服了传统线性防御体系只能应对单一性安全威胁的缺点。此外，我们还详细介绍了网络空间安全的基本范畴，并结合当前的热点应用，指出了立体式网络空间安全防御体系应采取的安全措施。

参考文献

[1] CLARK D D, WROCLAWSKI J, SOLLINS K R, et al. Tussle in Cyberspace: Defining Tomorrow's Internet [J]. IEEE/ACM Transactions on Networking, 2005, 13(3):462–475. doi: 10.1109/TNET.2005.850224
[2] 王晟, 虞红芳, 许都. 可信网络中安全、可控可管及可生存技术研究[J]. 中兴通讯技术, 2008, 14(1): 36–41
[3] NALEWAJSKI R F. Elements of Information Theory [J]. Perspectives in Electronic Structure Theory, 2011, 294(3–4): 371–395
[4] SCHWALM K T, SCHWALM K T. National Strategy to Secure Cyberspace [J]. Technical Report, AFRL–IF–RS–TR–2006–266, 2006, 1–27
[5] STEANE A M. How to Build a 300 bit, 1 Gop Quantum Computer [J]. Quantum Information & Computation, 2004, 7(3): 171–183
[6] 张阳, 陈开颜, 李雄伟,等. 基于差异度的密码芯片旁路攻击研究[J]. 通信学报, 2015, 36(03): 100–105
[7] 唐科萍, 许方恒, 沈才樑. 基于位置服务的研究综述[J]. 计算机应用研究, 2012, 12(12): 4432–4436
[8] SINGH B, DHAWAN S, ARORA A, et al. A View of Cloud Computing[J]. Communications of the ACM, 2013, 53(4): 50–58
[9] 朱洪波, 杨龙祥, 于全. 物联网的技术思想与应用策略研究[J]. 通信学报, 2013, 5(5): 31–31
[10] ZHANG Y Q, WANG X Y. A Symmetric Image Encryption Algorithm Based on Mixed Linear–Nonlinear Coupled map Lattice [J]. Information Sciences, 2014, 273: 329–351
[11] 郑东, 赵庆兰, 张应辉. 密码学综述[J]. 西安邮电大学学报, 2013, 18(6): 1–10
[12] MANATHA G S, SHARMA S C. Network Layer Attacks and Defense Mechanisms in MANETS– A Survey [J]. International Journal of Computer Applications, 2010, 27(1): 529–535
[13] ANDROULIDAKIS I. Mobile Phone Forensics [M]. Mobile Phone Security and Forensics. US: Springer, 2012: 75–99
[14] SHI Y, ZHENG Q, LIU J, et al. Directly

作者简介

张应辉，西安邮电大学通信与信息工程学院讲师、硕士生导师；主要研究方向为公钥密码学、云存储安全；目前主持国家自然科学基金；获得公开国家发明专利7项，其中授权2项，发表学术论文30余篇。

郑东，西安邮电大学通信与信息工程学院教授、博士生导师，西安邮电大学无线网络安全技术国家工程实验室主任；主要研究方向为基于编码的密码学、云存储安全；主持或参与了多项国家级研究课题，包括国家科技攻关项目、国家“863”计划项目等；出版学术专著2部，发表学术论文100余篇。

马春光，哈尔滨工程大学计算机学院教授、博士生导师；主要研究方向为信息安全与隐私保护、物联网等；主持完成了国家自然科学基金、教育部博士点基金等；获黑龙江省国防科技进步一等奖1项等，出版学术专著2部，发表学术论文60余篇。

DOI:10.3969/j.issn.1009-6868.2016.01.005
网络出版地址:http://www.cnki.net/kcms/detail/34.1228.TN.20151117.1512.010.html

面向数据的安全体系结构初步研究

Data Oriented Security Architecture

苗放/ MIAO Fang
(成都大学 大数据研究院,四川 成都 610106)
(The Research Institute of Big Data, Chengdu University, Chengdu 610106, China)

中图分类号:TN929.5 文献标志码:A 文章编号:1009-6868 (2016) 01-0019-004

摘要: 提出了一种以数据为核心和面向数据的信息安全解决方案,即面向数据的安全体系结构(DOSA)。DOSA将通过网络用户身份认证、数据所有权确立、数据注册、加密呈现、授权使用、水印记录、过程溯源、数据监管、安全应用等方法,建立一套开放环境下的数据安全体系,从数据的采集、管理、应用等层面上,最大限度地保护数据安全。在数据交易、数据存储、数据传输、数据应用、数据隐私保护等方面,具有较大的应用前景。

关键词: 面向数据的安全体系结构;数据所有权;加密呈现;数据安全

Abstract: System vulnerabilities and excess authority to access data are the main reasons for data divulgence. In this paper, a new data security solution called Data Oriented Security Architecture (DOSA) is proposed. DOSA is an open data security system to use a series of methods, including network user identity certification, data property rights, data register, data encryption present, data authority to use, data watermark log, data procedure tracing, data monitoring, and data safety application. DOSA protects data in data collecting, data management and data applications. It can be used in data transaction, data storage, data transfer, data application, data privacy safeguarding, etc.

Key words: data oriented security architecture; data property rights; encryption present; data security

1 信息安全面临的挑战

信息安全关乎国家安全、社会稳定、企业利益和个人隐私。随着环境的开放,数据的急剧扩张,人们对数据的依赖程度越来越高。由于数据集中存放、系统安全漏洞、数据越权访问等情况,使信息安全问题愈发突出。随着数据时代的到来,要求我们以新的数据体系结构去适应新的社会发展要求。

(1)开放环境下需要有新一代的数据安全解决方案

中国政府提出的"互联网+"行动计划,要将移动互联网、云计算、大数据、物联网等作为新时期经济发展的重要推力,信息系统或应用体系所面临的环境更为开放,对数据和信息安全的要求更高。

通常情况下,一个相对安全的信息系统或应用体系,是建立在一个相对封闭和安全的环境中,通过"门窗加固"等方式来保证这个封闭环境是安全的或可信的,更加强调的是网络空间安全、系统安全、环境安全和应用安全。虽然和外部交换信息时是通过数据加密或虚拟专用网络(VPN)通道来传输数据,但在这个相对"安全"的内部环境里,大多数数据却是处于"裸露"状态的。一旦有不速之客通过各种漏洞或非法获得权限进入到这个环境,裸露的数据就面临着极大的危险。

一些数据中心所涉及的数据安全,多是指利用数据备份、数据灾备等技术来保障数据不丢失、不被破坏,但仍存在着越权访问等危险行为,造成数据和信息泄露的隐患。

封闭环境下的安全方法在开放环境下的面临着极大的挑战,开放环境下的数据安全成为重要的研究课题。文献[1]提出了开放网络环境下数据的发布与管理,涉及了单向加密、新人凭证等属性概念;文献[2]较系统地了开放环境下敏感数据的安全问题;文献[3]提出了开放环境下敏感数据防泄露若干关键技术,其中主要涉及威胁模型、数据管理机制、可信执行、数据封装加密、秘钥保存等问题;文献[4]提出基于关系数据库上基于数据目的概念分层的隐私数据访问控制机制(R-PAACEE)模型的隐私分析算法,可以判断用户隐私,且该方法,对于结构化数据、日志数据、XML数据均有判断力,并且依据不同场景进行了实验分析。

在开放环境下,除了网络安全和系统安全保障之外,还需要在安全的

收稿日期:2015-11-10
网络出版时间:2015-11-17
基金项目:国家自然科学基金(61071121)

体系结构和安全的数据保护机制等方面有相应的举措。2012年，美国一些知名的数据管理领域的专家学者联合发布白皮书《Challenges and Opportunities with Big Data》，提出数据安全及系统架构问题的挑战[5]。

信息安全的核心就是数据的安全，开展面向数据和以数据为核心的数据安全体系研究是十分必要的。文献[6]在无线传感网络上的应用中对面向数据的安全模型进行过研究，但没有从面向数据的安全体系结构上开展研究。

因此，需要有一种新的安全体系结构，即面向数据的安全体系结构，来应对这个挑战。文章针对开放环境下数据安全性问题，进行了安全体系设计，引入了面向数据的技术架构，构建安全的数据访问机制。

（2）新时代下需要更底层的体系结构来保证数据安全和其应用

“互联网+”行动计划带给我们两点启示：一是以互联网为代表的信息技术集合由过去的行业性质，转变为了可以支撑其他行业发展的基础；二是只有“互联网+数据”，才能把传统行业加到互联网上去发展。

随着人类的发展，人们在地球上构建了不同的皮肤（见表1），让人类赖以生存、生活和发展，否则，就会被地球所淘汰。在由互联网构成的新皮肤上，承载着数据，使人的智慧得到提高，人类本身得到更好地发展。

人类经过漫长的文明发展之路，从物质文明进入到了非物质文明，亦即进入到目前以信息社会和数据时代为特征的非物质文明或文明3阶段，如表2所示。

由表2中可见，文明3的核心就是数据。从映射真实世界的虚拟世界，到信息、知识、智慧的根本，都是数据；从数据出发，才有信息、知识、智慧和决策；从网络连接传输的内容，到服务器、云主机、终端所存储、处理和展示的内容，也都是数据。数据是人类认识世界、沟通交流、获得知识、智慧决策的本源。一切技术、业务、功能、流程都是为了数据，并围绕数据而开展的。

▼表1 数据思维之地球皮肤概念

皮肤	皮肤类型	承载内容	作用
皮肤1	自然界	万物	世界上的生物通过阳光、空气、土地、水等生存
皮肤2	语言	人类	人类通过语言实现广泛交流
皮肤3	纸张	文字	文字通过纸张进行广泛传播和交流
皮肤4	交通	人类和物品	人类和物品通过陆地交通、水上交通、空中交通，快速到达地球上的任何地方
皮肤5	电力线	电	电通过电力线，给人们带来光明和动力
皮肤6	有线无线电波	电报、电话、广播、电视	电报、电话、广播、电视等信息内容，通过有线和无线电波，在更大范围、以更快速度传播信息到世界各地，从单向到双向
皮肤7	互联网	数据	各种以数据形式的媒体信息，通过互联网实现全球范围广泛和交互式地交流

▼表2 数据思维之人类文明演进轨迹

阶段	持续时间	社会	时代	技术手段	特征	征服事物
文明0	几万年	原始社会	蛮荒时代	自体能力	人类靠本能生存	蛮荒蒙昧，自食其力
文明1	几千年	农耕社会	庄园时代	农耕技术、畜牧技术	物质文明：人类开始征服生命物质	征服农作物和家畜家禽，利用生物为人类服务
文明2	几百年	工商社会	帝国时代	建筑技术、冶炼技术、机械技术、制造技术、电力技术等	物质文明：人类开始征服无生命物质	征服煤、石油、金属、非金属，利用非生物为人类服务
文明3	几十年	信息社会	数据时代	计算机、互联网、物联网、社交网、云计算、智能技术、大数据	非物质文明：人类开始征服思想、智慧，核心为数据	征服人类自己的思想世界、精神世界、意识世界，利用信息和数据为人类服务

数据在文明3和非物质文明中是至关重要的基本要素，因此以数据视角来看待文明3，就需要有面向数据的体系结构和安全体系结构，来支撑非物质文明的社会发展。为此，作者提出面向数据的体系结构（DOA）[7]，来构建数据时代的底层架构，并试图去解决数据所有权、信息共享、系统功能扩展、数据管理、大数据分析和挖掘支持、软件工程、信息安全、数据拥有者利益保障等问题。

2 面向数据的安全体系结构

面向数据的安全体系结构（DOSA）旨在从架构角度对未来的数据安全体系进行全方位设计，包括数据的管理和应用等。DOSA是在DOA基础之上，面向数据和以数据为核心的关于数据的安全体系结构，构建起从数据保护到授权应用的整套机制。

DOSA建立在云计算基础之上，以数据“天生加密、授权使用”为原则，对数据的属性进行注册和管理，实现数据的安全管理和安全的相关应用。

中国颁布的《电子签名法》，从法律和技术层面上，为面向数据的安全体系结构奠定了重要基础。《电子签名法》所依赖的用户认证中心（CA）和公共密钥基础设施（PKI）技术，是面向数据的安全体系结构的基本数学和技术保障。

作为非物质社会的基本元素，数据应满足以下的基本特征：具有广义数据的概念，并有生命和属性（具有身份属性、安全属性、时间以及空间属性）。

（1）广义数据：凡是能够被计算机注册和登记的任何事物都称之为数据。

（2）身份属性：数据权属，即数据

的主人(数据生产者和数据所有者)、朋友(数据使用者或被授权人)、陌生人(未授权和待授权人)和敌人(不授权人、黑名单)。

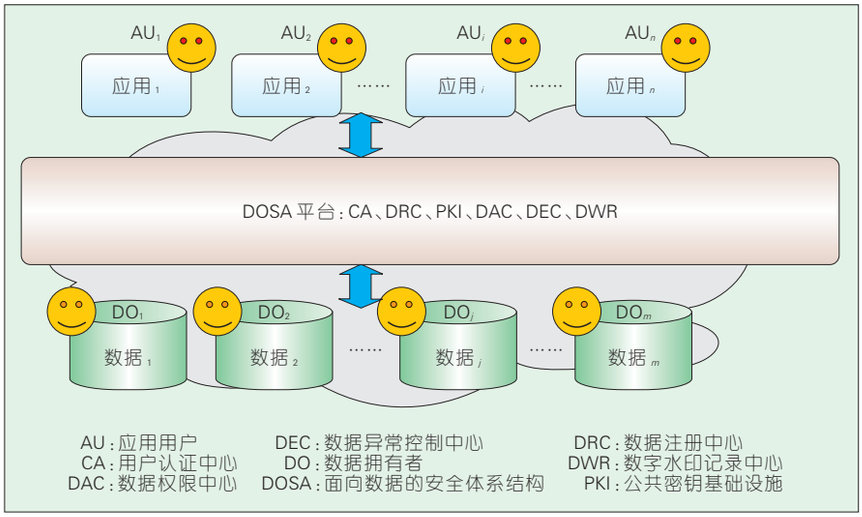
(3)安全属性:数据具有自保护功能,要“穿戴盔甲”,以加密方式呈现,具有不同的加密级别和深度,数据的使用要经过授权。

(4)时间和空间属性:数据的产生、授权以及使用等,都有时间和空间印记。

数据是应用的基础,不依赖于特定的硬件环境和软件环境,同一数据可以支撑不同的应用。

为便于管理,我们将数据分成存储和传输时保持加密的“数据”态和在应用中授权使用时解密的“应用态”。数据只有在应用态时是处于解密状态,一旦完成应用或离开了应用环境,或是由应用产生了新的数据,数据应立即变为加密的数据态,充分保证数据的安全及使用的授权。数据态的数据,既适合于封闭环境,也适合于较为开放环境,而应用态的数据,仅适合于相对来说比较封闭的环境[8-10]。

DOSA由以下主要部件构成:CA、数据注册中心(DRC)、PKI、数据权限中心(DAC)、数据异常控制中心(DEC)、数字水印记录中心(DWR)以及数据应用单元(DAUs)等,来构成面向数据的安全体系结构,从数据管理、数据安全保障到安全应用的全过程管理,如图1所示。



▲图1 DOSA平台与数据和应用之关系

2.1 CA用户认证

DOSA的一个核心理念是要确定数据与用户的关系,需要明确数据的所有权人。这就需要对参与网络活动的用户进行注册和身份确认。DRC要对所有用户进行登记注册,而用户身份则通过CA来进行认证。CA认证采用第三方CA认证中心,对网络用户颁发数字证书,即公钥和私钥。私钥以多种形式安全地发放到每位用户的手中,而公钥则存储在数据注册中心中。

2.2 DRC数据注册

DRC是DOSA的核心部件,用于注册各种数据的属性信息,包括数据的安全属性信息和数据权人信息等,并对数据的使用过程进行记录。DRC还要保留所有数据权人和应用用户的公钥。

DRC用来构建逻辑的数据资源池,通过建立索引和搜索引擎,实现数据和应用的管理和服务。一个DRC可以关联其他的DRC,从而实现广泛的数据共享。

2.3 PKI数据所有权与加密呈现

一旦数据产生了,DOSA平台就需要明确两件事情:一是要确定数据的生产者和数据的所有者;二是要对数据进行加密,防止他人窃取。一般情况下,数据的生产者就是数据的所有者。但在有些情况下,这两者是不一样的:数据的生产者不一定是数据的所有者。

确定数据的生产者,需要用数据生产者的私钥对数据进行加密(数字签名),标明数据的生产者身份。确定数据的所有者,则需要要用所有者的公钥加密来确定,同时实现数据的天生加密。对于体量较大的数据,可以采用对称密钥加密的办法,对对称密钥再进行公钥加密。不论数据是处于存储状态还是传输状态,都要保持加密呈现。

换句话说,用公钥加密,就确定了数据的所有者,并做到了数据加密呈现;用私钥加密,就确定了数据的生产者,数据生产者一般情况下就是数据所有者,但有些情况下不同。

2.4 DAC数据授权使用

DAC也是DOSA的关键部件,用于对数据进行授权管理。数据在生成、存储和传输时是加了密且不可使用的,而经过授权的用户在使用数据时才是解密和可访问的。数据授权,就是数据的权属变更,就是数据解密和加密的过程,即用数据所有者的私钥解密后再用数据使用者(被授权人)的公钥加密,授权过程[11]要通过水印和数据注册中心进行记录和管理。对于体量较大的数据,采取的是对称密钥加密方法,授权过程只是对对称密钥进行。

DOSA下的数据安全使用、记录,网络用户的认证等,可采用网络安全的验证、授权和记账(AAA)等技术。

2.5 DEC数据监管

DEC是DOSA的重要部件,用于

对数据资源进行自适应管理，保证数据的唯一性和一致性，监管和处置数据的各种异常行为。

2.6 DWR数字水印记录

DWR以水印的方式将数据所有者及授权使用过程记录下来，与原始数据一起进行加密管理，便于数据的溯源、记账和数据的非授权使用的取证。

2.7 DAUs数据安全应用

DAUs用于关联应用对数据的访问，并对各种应用提供支持。要确定数据安全应用的环境，一般考虑数据在内存中解密使用，要通过多种手段实现内存数据的安全保障和不被侵入窃取。

3 DOSA应用展望

DOSA作为一种数据安全理念和机制，就是要保证数据能够在数据和应用两个层面中都能做到安全、可靠，便于管理和使用，既可以在传统的封闭环境下应用，增强数据的安全保护，又可以在开放环境下保护数据的安全和不被越权访问[12-13]。

目前有关信息安全、数据安全的理论和方法体系，有关网络安全的AAA技术，有关CA技术、PKI技术、密钥体系、加解密技术，有关可信技术，以及不断发展的网络空间安全技术、系统安全技术、应用环境安全技术等，都能在DOSA框架下使用，但需要进一步从面向数据和以数据为核心的角度，进行重新梳理，从数据安全的理念、理论、方法和受保护数据的应用机制等方面，进行适应性和深入地研究，为进一步提高信息安全提供保障[14-15]。

基于DOSA，目前正在试点开展以下一些应用：

（1）数据交易（虚拟数字资产保护及交易）平台。在建立数据资产所有权的基础上，通过数据加密呈现、授权交易、过程记录、价值评估、记账计费管理、水印溯源等，保障数据安全交易和数据拥有者利益。

（2）数据隐私保护。通过分析数据和隐私的特征，进行数据脱敏、数据所有权确认、数据加密、数据授权应用、数据安全应用、数据过程记录和溯源等，进行一些数据的隐私保护研究。

4 结束语

开放环境下信息安全问题集中体现在数据的安全上。DOSA采用面向数据和以数据为核心的理念，建立数据与用户之间的权属关系，采用数据“天生加密，授权使用”方法，通过CA、DRC、DAC、DEC、PKI、DWR、DAUs等实现数据的安全管理和安全应用，建立从数据保护到授权应用的整套机制。

基于DOSA的初步应用表明：面向数据的安全体系结构能有效解决和应对开放环境下数据的安全、数据所有权、数据交易、数据共享、数据管理、数据隐私保护等问题和挑战。

致谢

本研究得到中国软件行业协会赵小凡理事长、北京邮电大学杨义先教授、四川省计算机学会宋昌元秘书长以及成都大学大数据研究院、成都理工大学空间信息技术研究所、成都灵云信息技术有限公司、成都五舟汉盛科技有限公司、四川红山世纪科技有限公司等的大力支持和帮助，谨致谢意！

参考文献

[1] 朱静波. 网络环境下敏感数据的发布和管理[D]. 杭州: 浙江大学, 2006
[2] 陈珂. 开放式环境下敏感数据安全的关键技术研究[D]. 杭州: 浙江大学, 2007
[3] 闫玺玺. 开放网络环境下敏感数据安全与防泄密关键技术研究[D]. 北京: 北京邮电大学, 2012
[4] 刘逸敏. 基于访问目的的隐私数据访问控制机制研究[D]. 上海: 复旦大学, 2012
[5] AGRAWAL D, BERNSTEIN P, BERTINO E, et al. Challenges and Opportunities with Big Data–A Community White Paper Developed by Leading Researchers Across the United States [EB/OL].[2012–10–02]. http://www.cra.org/ccc/files/docs/init/bigdatawhitepaper.pdf
[6] XIAO X R, SUN X M, WANG X B, et al. DOSM: A Data–Oriented Security Model Based on Information Hiding in WSNs [J]. Information Technology Journal, 2009, 8(5): 678–687
[7] 苗放.DOA（面向数据的体系结构）[EB/OL]. http://baike.baidu.com/subview/649092/12822804.htm#viewPageContent, 2014
[8] 尹建国. 美国网络信息安全治理机制及其对我国之启示[J]. 法商研究, 2013, (2): 138–146
[9] 赵勇. 大数据革命:理论、模式与技术创新[M]. 北京:电子工业出版社, 2014
[10] 王世伟. 论信息安全、网络安全、网络空间安全[J]. 中国图书馆学报, 2015, 41(2): 72–84. doi:10.13530/j.cnki.jlis.150009
[11] KRESIMIR S, HRVOJE O, MARIN G. The Information Systems' Security Level Assessment Model Based on An Ontology and Evidential Reasoning Approach [J]. Computers & Security, 2015, 55(6): 100–112.doi: doi:10.1016/j.cose.2015.08.004
[12] GEORGE S, VLADLENA B, JEAN N E, et al. Individual Information Security, User Behaviour and Cyber Victimisation: An Empirical Study of Social Networking Users [J]. Technological Forecasting and Social Change, 2015, (5): 320–330. doi:10.1016/j.techfore.2015.08.012
[13] GURPREET D, ROMILLA S, CRISTIANE P. Interpreting Information Security Culture: An Organizational Transformation Case Study [J]. Computers & Security, 2015, (4): 63–69.doi: 10.1016/j.cose.2015.10.001
[14] 程学旗, 靳小龙, 王元卓, 等. 大数据系统和分析技术综述[J]. 软件学报, 2014, 25(9): 1889–1908
[15] 孟小峰, 慈祥. 大数据管理:概念、技术与挑战[J]. 计算机研究与发展, 2013, (1): 146–169

作者简介

苗放，成都大学大数据研究院院长，成都理工大学空间信息研究所所长、教授、博士生导师；主要研究方向为空间信息技术及应用、大数据管理；先后主持国家自然科学基金项目、国家“863”、“973”计划子课题、部省项目20余项，成果获得部省级二等奖2项，三等奖2项；已发表论文160余篇，其中被EI检索20余篇。

DOI:10.3969/j.issn.1009-6868.2016.01.006
http://www.cnki.net/kcms/detail/34.1228.TN.20151225.0953.002.html

安全多方计算技术研究与应用

Research and Application of Secure Multi-Party Computation

张卷美/ZHANG Juanmei
徐荣华/XU Ronghua
(北京电子科技学院,北京 100070)
(Beijing Electronic Technology Institute, Beijing 100070, China)

中图分类号:TN929.5 文献标志码:A 文章编号:1009-6868 (2016) 01-0023-003

摘要: 认为同态密码的本质是通过密文运算,实现相对应的明文运算。基于同态密码、格理论密码,分别设计了安全多方计算协议,解决了安全两方线段求解直线相交问题和聚类分析中一种经常遇到的加权平均问题。认为目前安全多方计算的实际应用比较滞后,但随着其理论的不断成熟以及各种密码理论基础技术的不断发展,安全多方计算最终会为新时代下的信息安全提供服务。

关键词: 安全多方计算;同态加密;格密码

Abstract: The essence of homomorphic encryption is to realize the corresponding plaintext operation by calculating cipher text. In this paper, we propose some secure multi-party computation schemes based on homomorphic encryption and lattice theory. With these protocols, the secure two-party line segment intersection problem and weighted-average problem, which are often encountered when solving the problem of clustering analysis, are solved. Practical application of secure multi-party computation is lagging, but with the continuous development of its theory and various kinds of cryptography, secure multi-party computation will increase information security in the future.

Key words: secure multi-party computation; homomorphic encryption; lattice homomorphism

安全多方计算是密码学的基础问题之一,概括了大多数密码协议,如认证协议、在线支付协议、公平交换协议、拍卖协议、选举协议、密文数据库查询与统计等等。在电子选举、电子投票、秘密共享等场景有关广泛的应用[1-2]。

1 安全多方计算的概念

针对安全多方计算,2004年Goldreich给出了一个简单、完整、统一形式化定义[3]。他将安全多方计算抽象成一个随机过程:U_1、U_2、……、U_n是n互不信任的参与方,共同计算函数$f(x_1,x_2,\cdots,x_n)=(y_1,y_2,\cdots,y_n)$,其中函数$f$是一个计算复杂度为概率多项式的随机函数,$U_i$拥有秘密输入$x_i\in(0,1)^*$,通过计算希望获得$y_i\in(0,1)^*$,但不向其他参与方泄露任何信息。

在理想的世界中,假设第三方是可信的,计算函数f,则其计算时间亦为概率多项式时间,抽象为交互式图灵机,其执行过程为可信第三方收集各方输入,然后计算结果,再分别秘密发送结果给对应参与方。理想世界中敌手IA执行协议过程中获取到的输出变量为$IDEAL(\mathrm{IA})$。

在现实的世界中,可信第三方不存在,同样计算函数f,则需要收集各方输入,然后计算输出结果,再分别秘密发送结果给对应参与方。理想世界不同的是敌手RA可以窃听并收集诚实参与方之间的通信信息,但不能修改其通信内容。现实世界中敌手IA执行协议过程中获取到的输出变量为$REAL(\mathrm{RA})$。

对于现实世界中的任意敌手RA,都存在相应理想敌手IS,若在计算过程中,使得$IDEAL(\mathrm{IA})$和$REAL(\mathrm{RA})$计算不可区分,即$IDEAL(\mathrm{IA})\approx REAL(\mathrm{RA})$,则认为安全多方计算协议是安全的。

2 同态加密理论在安全多方计算中的应用

同态加密[4]能够在不对密文解密的情况下,对密文进行计算,从而实现对明文的计算,这与安全多方计算中在不泄露任何数据隐私信息的情况下完成安全计算的需求不谋而合。

目前,同态密码是密码学领域研究的热点之一。同态的分类较为常见的是:单同态、双同态、无限同态和有限同态,这是一个较为概括性的分类方法。

(1)单同态是指关于明文的某一种运算具有同态特性的同态,分为乘法同态和加法同态。

收稿日期:2015-11-20
网络出版时间:2015-12-25

• 乘法同态

对于加密体制 $ALG(E,D,P,C)$，$x\in P$，$y\in P$，如果满足 $D(E(x)@E(y))=xy$（@为密文空间 C 的操作），则称 ALG 满足乘法同态。

• 加法同态

对于加密体制 $ALG(E,D,P,C)$，$x\in P$，$y\in P$，如果满足 $D(E(x)\#E(y))=xy$（#为密文空间 C 的操作），则该体制 ALG 满足加法同态。

（2）双同态指关于明文空间的加法运算和乘法运算都是同态的，且明文空间必须是一个环。

（3）全同态指关于明文空间可以实现任何运算的同态，即对明文空间的任何运算都可以转化为密文空间恰当的运算解密值。无限的双同态密码体制，可以转化为全同态。

通过对现有的具有同态性质的加密体制进行分析，我们得出：原始的ElGamal加密体制满足乘法同态、RSA满足加法同态、Paillier满足加法同态等。虽然目前没有成熟易用的同态密码体制能够满足任意形式的计算需求，但是已经存在的成熟且具有单一同态性质的密码算法就能够满足部分安全多方计算场景中的应用[5-6]。

利用具有加法同态密码体制可以求解百万富翁问题。两个富翁分别为A和B，并且有一个满足加密的同态加密算法 $ALG(E,D,P,C)$，假设A的财富为 m_1，B的财富为 m_2，分别用加密算法对财富进行加密：$E(m_1)=M_1$，$E(m_2)=M_2$，得出 (M_1,M_2) 的大小就可以得到 (m_1,m_2) 的大小。因为 $D(E(m_1)-E(m_2))=m_1-m_2$。

我们利用ElGamal密码体制可以求解两私有点的直线方程问题，实际上就是要秘密求出两私有点坐标差商的问题。这就是著名的安全两方线段求交问题。

在求解之前，我们先设计一个新的ElGamal密码体制。

（1）系统参数为：选择一个大的素数 p，g 是循环群 Z_p^* 的生成元，再选一个随机数 $x\in Z_p^*$，计算 $y=g^x \bmod p$，私钥为 x，公钥为 (y,g,p)。

（2）加密过程为：对任意的消息 m，选随机数 k，满足 $\gcd(p-1)=1$，计算密文 $E(m)=(a,b)=(g^k \bmod p, y^k g^m \bmod p)$。

（3）解密过程为：计算 $\log_g(b/(a)^x)\bmod p$，得到明文。

显然，求对数要付出很大的计算代价，需要在 Z_p^* 的空间里搜索结果。本方案用预先计算，查表解密的方法实现。上述密码算法有加法同态特性。假设 (m_1,m_2) 为两个消息：

$E(m_1)=(a_1,b_1)=(g^{k_1} \bmod p, y^{k_1} g^{m_1} \bmod p)$，

$E(m_2)=(a_2,b_2)=(g^{k_2} \bmod p, y^{k_2} g^{m_2} \bmod p)$，

在密文空间定义一种预算 $E(m_1)@E(m_2)=(a_1a_2 \bmod p, b_1b_2 \bmod p)$，则有：

$$E(m_1)@E(m_2)=(g^{k_1+k_2} \bmod p, y^{k_1+k_2} g^{m_1+m_2} \bmod p) \quad (1)$$

显然，$D(E(m_1)@E(m_2))=m_1+m_2 \bmod p$ 满足加法同态特性。

上述密码体制满足与常数 n 的乘法同态性下文称为常数乘法同态，也可以理解为一种特殊的求 n 次和的加法同态，则有：

$$\begin{aligned} D(E(m)^n) &= D(E(m)@E(m)@\cdots@E(m)) \\ &= D(g^{nk} \bmod p, y^{nk} g^{nm}) \bmod p \\ &= nm \bmod p \end{aligned} \quad (2)$$

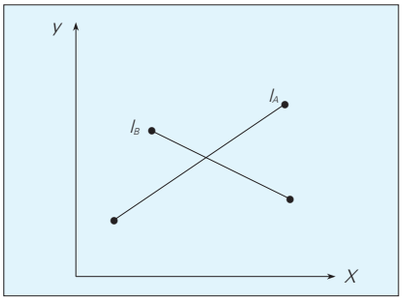
在安全两方线段求交问题中，假设Alice拥有线段 $l_A: y=a_1x+b_1 \bmod p$，Bob拥有线段 $l_B: y=a_2x+b_2 \bmod p$。Alice和Bob希望计算两条线段的交点。计算结束后，除了交点的坐标信息外，对方不能获知其他任何信息，如图1所示。

上述问题的实质就是求解式(3)。

$$\begin{cases} y=a_1x+b_1 \\ y=a_2x+b_2 \end{cases}$$ 其中：

$$m_1\leqslant x\leqslant n_1,\ m_2\leqslant x\leqslant n_2 \quad (3)$$

当 $a_1-a_2\neq 0$ 时，解为 $x=\dfrac{b_2-b_1}{a_1-a_2}$，$y=a_1x+b_1$



▲图1 安全两方线段求解直线相交问题

安全多方计算协议[7]设计为：Alice拥有 (a_1,b_1)，对应密文为 $A(E(a_1),E(b_1))$；Bob拥有线段 (a_2,b_2)，对应的密文为 $B(E(a_2),E(b_2))$。

利用上述新设计的密码体制，通过密文计算可以得 $x=\dfrac{r(b_2-b_1)}{r(a_1-a_2)}=\dfrac{b_2-b_1}{a_1-a_2}$，并可得出安全两方线段的交点：$\left(\dfrac{b_2-b_1}{a_1-a_2}, a_1x+b_1\right)$。

类似于这样的问题还有很多，如判断3个私有点共线问题等，这些问题都可以采用该协议来解决。

3 格理论在安全多方计算中的应用

格理论的研究源自于1611年Kepler所提出的猜想。经过400多年的发展，格的困难问题依次被提出来，最后确定了格上的主要困难问题：最短向量问题（SVP）、最近向量问题（CVP）、小整数解问题（SIS）、错误学习问题（LWE）等。早期的格理论主要应用于密码分析，如1982年Lenstra等提出的LLL格基规约算法，该算法能够有效求解出近似于最短向量的格基。直到1996年，Ajtai证明了在最坏情况下与平均情况下求解格困难问题是等价的，并给了设计格密码方案新的想法，这对于研究格密码方案具有重要的意义。

格理论上的安全多方计算协议是基于格公钥密码体制，NTRU公钥密码体制是一种基于格的公钥密码体制。

（1）系统参数为：3个公开参数为

(N,p,q)，通常情况下 $p=3,q=2^k$，$N-1$ 是多项式的最高次数，$*$ 表示卷积乘，设 $a(x),b(x)\in R$，则 $c(x)=a(x)*b(x)=\sum_{k=0}^{N-1}[\sum_{i+j=k \bmod N} a_i b_j]x^k$。它构建在商环 $Z[x]/(x^N-1)$ 上。$L(a,b)$ 表示环中具有 a 个系数为1，b 个系数为-1，其余系数均为0的全体整系数多项式。随机选取两个多项式 $f=1+pF$ 和 $g\in L(d_g,d_g)$，其中保证 f 存在逆元 f_p 和 f_q，使得 $f\times f_p=1(\mathrm{mod}\,p)$，$f\times f_q=1(\mathrm{mod}\,q)$。计算 $h=f_q\times g(\mathrm{mod}\,q)$，NTRU的公钥为 (N,p,q,h)，私钥为 f。

（2）加密过程为：用户选取随机多项式 $r\in L(d_r,d_r)$，对于明文消息 m，计算 $c=pr*h+m(\mathrm{mod}\,q)$ 得到密文。

（3）解密过程为：解密者得到密文 $c=pr\times h+m(\mathrm{mod}\,q)$ 后，首先计算 $a\equiv c\times f(\mathrm{mod}\,q)$，再计算 $m'=a\times f_p$，最后计算 $m\equiv m'(\mathrm{mod}\,p)$。

因为 $E(x)+E(z)=p(r_1+r_2)h+(x+z)(\mathrm{mod}\,q)$，令 $r_5=r_1+r_2$，则可得出：

$$E(x)+E(z)=pr_5h+(x+z)(\mathrm{mod}\,q)=E(x+z) \quad (4)$$

因此NTRU公钥加密体制满足加法同态性质。

又因为 $E(x)=pr_1h+x(\mathrm{mod}\,q)$，$z\times E(x)=pzr_2h+zx(\mathrm{mod}\,q)$，令 $r_6=x*r_2$，则可得出：

$$z\times E_A(x)=E_A(zx) \quad (5)$$

因此，可以认为NTRU公钥加密体制满足乘法混合同态性质。

聚类分析中有一种常用遇到的是加权平均问题（WAP），该问题描述为：A拥有 (x,m)，其中 x 为一实数，m 是一个整数；B拥有 (y,n)，希望能够联合计算 $\frac{x+y}{m+n}$。在加权平均问题中，整数 n 和 m 需要被保护，双方需要在不知对方任何消息的条件下，联合完成计算，并且需要保证计算结果的正确性。

上述加权平均问题已在文献[8]中给予解决。文中假设有 n 个用户 $(P_1,P_2,\cdots,P_n)$，每个用户拥有 (x_i,y_i)（其中 $i=1,2,\cdots,n$），安全计算为 $\frac{x_1+x_2+\cdots+x_n}{y_1+y_2+\cdots+y_n}$。

假设参与安全多方计算的用户分别为 $P_1,P_2,\cdots,P_n$，选取 P_n 作为最终结果的计算者，基于NTRU设计的多方安全计算协议流程如下：

（1）P_n 选取可逆多项式 z，通过采用文献[9]两方安全计算协议，可从 P_1 处得到 $z(x_1+x_n)$ 和 $z(y_1+y_n)$，从 P_i 处得到 $z(x_i+x_n)$ 和 $z(y_i+y_n)$，$(i=1,2,\cdots,n)$。

（2）结合收集到的数据，P_n 采用NTRU体制的加法同态性将所有数据相加，计算得到 $z(x_1+x_2+\cdots+n\times x_n)$ 和 $z(y_1+y_2+\cdots+n\times y_n)$。

（3）计算 $z(x_1+x_2+\cdots+n\times x_n)-(n-1)\times z\times x_n=z(x_1+x_2+\cdots+x_n)$，采用同样的方法得到 $z(y_1+y_2+\cdots+y_n)$。

（4）计算得到 $\frac{x_1+x_2+\cdots+x_n}{y_1+y_2+\cdots+y_n}$。

（5）将 $E_{P_i}(\frac{x_1+x_2+\cdots+x_n}{y_1+y_2+\cdots+y_n})$ 发送给用户 P_i，其中 $i=i,2,\cdots,n$。

（6）各个用户通过解密可以获得将 $\frac{x_1+x_2+\cdots+x_n}{y_1+y_2+\cdots+y_n}$。

上述安全多方计算协议借助交互多方中的某一位成员进行计算，其他成员只与该成员进行交互，最终计算方将所得结果秘密传送给各个成员。该方案类似于单方交互协议，可以最终实现安全多方计算，但是在效率上有待提高。

4 结束语

作为密码学研究的一个重要方向，安全多方计算既古老又年轻。与理论研究成果相比，安全多方计算的实际应用比较滞后，主要是效率和安全性还不能完全满足现实的需求。根据安全多方计算的特点，我们已经为各种不同的应用场景设计了相应的安全多方计算方案，主要集中在大数据隐私保护、云计算和文数据库检索和统计等安全性需求较高的应用场景中。未来还有很多研究要做，主要集中在恶意型下的多方安全计算问题中，因为恶意模型环境下参与方可以完全不遵守协议规则。虽然在现阶段，安全多方计算的应用还存在困难，但是随着安全多方计算理论的不断成熟以及各种密码理论基础技术的不断应用，安全多方计算最终会走入我们的实际生活，为互联网时代、云计算时代、大数据时代的信息安全服务。

参考文献

[1] YAO Q Z. Protocols for Secure Computations [C]// Proceedings of 23rd Annual IEEE Symposium on Foundations of Computer Science，Chicago, USA, 1982: 160– 164
[2] GOLDREICH O, MICALI S, WIGDERSON A. How to Play Any Mental Game or A Completeness Theorem for Protocols with Honest Majority [C]//STOC 1987, 1987: 218–229
[3] GOLDREICH O. Foundations of Cryptography: Volume II – Basic Applications [M]. Britain: Cambridge University Press, 2004
[4] GENTRY C. A Fully Homomorphic Encryption Scheme [D]. USA: Stanford University, 2009
[5] DU W L. Secure Multiparty Computation Problems and Their Applications[C]//A Review and Open Problems New Security Paradigms Workshop 2001, Cloudcroft , New Mexico, USA, 2001
[6] 刘文, 王永滨. 安全多方信息比较相等协议及其应用[J]. 电子学报, 2012, 40(5): 871–876
[7] 陈志伟，张卷美，李子臣. 基于ElGamal变体同态的安全两方计算协议设计[J]. 通信学报, 2015, 36(2): 1–8
[8] 刘立强，李子臣.一种基于NTRU的安全两方计算协议[C]//全国信息隐藏暨多媒体信息安全学术大会（CIHW），北京, 中国, 2012
[9] 胡予濮. 一个新型的NTRU类数字签名方案[J]. 计算机学报, 2008, 31(9): 1661–1666

作者简介

张卷美，北京电子科技学院基础部副教授；从事计算数学的教学和研究工作；主持参加省级学术研究、教改项目6项；发表学术论文和教改论文20余篇，编写出版教材5部。

徐荣华，北京电子科技学院基础部教师，长期从事代数、密码学教学与科研等工作。

DOI:10.3969/j.issn.1009-6868.2016.01.007
网络出版地址:http://www.cnki.net/kcms/detail/34.1228.TN.20151117.1520.016.html

同态加密的发展及应用

The Development and Applications of Homomorphic Encryption

巩林明/GONG Linming
李顺东/LI Shundong
郭奕旻/GUO Yimin
(陕西师范大学计算机科学学院,陕西西安 710062)
(School of Computer Science, Shanxi Normal University, Xi'an 710062, China)

中图分类号:TP393 文献标志码:A 文章编号:1009-6868 (2016) 01-0026-004

摘要: 认为密码学中的同态加密技术可以为分布式计算环境的用户隐私保护提供强有力的技术支撑。同态加密方案被分成3种类型:部分同态加密、浅同态加密和全同态加密。同态加密方案在分布式计算环境下的密文数据计算方面有着重要的应用,包括:安全云计算与委托计算、远程文件存储、密文检索等。指出目前全同态加密方案的构造还处于理论阶段,尚不能用于实际的密态数据计算问题,如何设计基于代数系统的(自然)全同态加密方案依然是未来研究的重点。

关键词: 同态加密;密态计算;安全多方计算;安全云计算;分布式计算

Abstract: Homomorphic encryption has been widely used to provide data security and privacy for users in a distributed computing environment. There are three types of homomorphic encryption schemes: part homomorphism, somewhat homomorphism and fully homomorphism. Homomorphic encryption schemes have many important applications in computing based on ciphertext, including secure cloud computing and outsourcing, remote file storage, and search on encrypted data. The constructions of the fully homomorphic encryption scheme is still in the theoretical stage and can not be used for actual data calculation. How to develop the nature-fully homomorphic encryption schemes based on algebra is still the focus in the future research.

Key words: homomorphic encryption; privacy computing; secure multi-party computing; secure cloud computing; distributed computing

Rivest、Adleman和Dertouzos[1]于1978年提出了秘密同态的思想:对几个数据的加密结果进行运算后再解密,得到的结果与这些数据未加密时执行某一运算所得的结果一致。此后,研究人员在同态加密方案设计方面做了大量的工作并取得了大量的研究成果。例如,1978年由Rivest、Adleman和Dertouzos[2]提出的RSA加密系统、1985年由ElGmal提出的ElGmal加密方案[3]、1998年由Okamoto和Uchiyama[4]提出的《A new public-key cryptosystem as secure as factoring》、1999年由Paillier[5]提出的Paillier加密方案、2002年由Domingo-Ferrer提出的《A provably secure additive and multiplicative privacy homomorphism》、2005年由Boneh [6]等提出的用于保密计算2-析取范式(2DNF)的加密方案、2009年由Gentry等[7]首次提出的全同态加密(FHE)方案、2010年Dijk[8]等提出的整数域上的FHE方案、2011年由Brakerski、Vaikuntanathan[9]两人提出的基于误差学习的FHE方案、2012年由Brakerski和Gentry[10]等提出的无需电路自举的分层FHE方案、同年由Brakerski[11]提出的无需模转换的FHE方案、2013年由Gentry[12]等提出的环上的FHE方案、2014年由Brakerski[13]等提出的基于标准误差学习的FHE方案、2015年由Cheon[14]等提出的基于中国剩余定理的整数域上的FHE方案等都是比较著名的同态加密成果。

目前出现的同态加密方案可被分成3种类型:部分同态加密、浅同态加密和全同态加密。部分同态只能实现某一种代数运算(或、乘、加);浅同态能同时实现有限次的加运算和乘运算;全同态能实现任意次的加运算和乘运算。

同态加密方案,除了可以实现加密功能外,还可以用于密文数据的计算。近些年来随着网络技术的发展,以同态加密技术为支撑的密文数据计算越来越多地被应用于各种分布式计算中,例如,安全云计算与安全云存储中有关用户隐私的保护和高效的安全多方计算协议,都需要同态加密技术支持;其他应用如电子选举、远程文件存储、密文检索、版权保护等也都需要同态技术的支持。

收稿日期:2015-11-08
网络出版时间:2015-11-17

1 同态加密系统中的一些定义

(1)同态性

假设一个加密系统的加密函数与解密函数分别为 $E:\mathcal{M}\rightarrow C$ 与

$D:C\to\mathcal{M}$，其中 $\mathcal{M}$ 与 C 分别为明文空间与密文空间；令 $\boxplus$ 和 $\oslash$ 分别为定义在明文空间和密文空间上的代数运算或算术运算。则加密方案的同态性定义为：给定任意的两个 $m_1,m_2\in\mathcal{M}$，如果一个加密系统的加密函数与解密函数满足代数关系 $m_1\boxplus m_2=D(E(m_1)\oslash E(m_2))$ （或 $E(m_1\boxplus m_2)=E(m_1)\oslash E(m_2)$），则称该加密系统具有同态性。

（2）加法、乘法、异或同态

一个具有同态性的加密系统，若明文空间上的运算为代数加法“+”，则该加密系统被称为加法同态加密系统；若明文空间上的运算为代数乘法“*”，则该加密系统被称为乘法同态加密系统；若明文空间上的运算为算数异或“$\oplus$”，则该加密系统被称为异或同态加密系统。

（3）浅同态与全同态

只满足一种代数（或算术）同态运算的加密系统，被称作部分同态加密系统；同时满足加法和乘法同态运算，且只能进行有限次乘法或加法运算的加密系统称为浅同态加密系统；同时满足加法和乘法同态的加密系统称为全同态加密系统。

2 同态加密的发展

同态加密思想从提出到现在，在具体实现方案方面，经历了3个重要时期：1978—1999年是部分同态加密的繁荣发展时期；1996—2009年是部分同态加密与浅同态加密的交织发展时期，也是浅同态加密方案的繁荣发展时期；2009年以后是全同态加密的繁荣发展时期。下面将以时间为主线，按照同态加密方案的类型介绍同态加密的发展。

2.1 部分同态加密方案

部分同态加密方案按照明文空间上能实现的代数或算术运算分为乘法同态、加同态和异或同态3种类型。下面从几个著名的同态加密方案的优缺点入手，总结一下乘法同态、加同态、或同态加密方案的特性。

（1）乘法同态加密方案。乘法同态加密方案的同态性表现为 $m_1\times m_2=D(E(m_1)\times E(m_2))$。RSA[5]是最早的具有乘法同态性的加密方案，它是基于因子分解困难问题的，属于确定性加密，不能抵御选择明文攻击；1985年，ElGamal[3]基于有限域上的离散对数困难假设设计了ElGamal加密算法，该加密方案同样具有乘法同态性，并且满足选择明文不可区分（IND-CPA）安全。

（2）加法同态加密方案。加法同态加密方案的同态性表现为 $m_1+m_2=D(E(m_1)\oslash E(m_2))$（$\oslash$为定义在密文空间上的某种代数运算或算术运算）。具有加法同态性的加密方案有很多，应用最为广泛的当属Paillier[5]加密系统，该加密系统基于高阶合数度剩余类困难问题，且具有IND-CPA安全。

（3）异或同态加密方案。乘法同态加密方案的同态性表现为 $m_1\oplus m_2=D(E(m_1)\oslash E(m_2))$（$\oslash$为定义在密文空间上的某种代数运算或算术运算）。目前，只有Goldwasser-Micali[15]加密系统属于该类同态加密系统，该加密系统基于二次剩余困难问题，虽具有IND-CPA安全，但每次只能加密单比特，因此加密效率会比较低。

2.2 浅同态加密方案

浅同态加密方案能同时进行有限次乘法和加法运算的加密。从某种程度上讲，该类型的加密方案是人们在研究解决RSA 3个人提出的公开问题（如何设计全同态加密方案）的过程中，出现的“副产品”。1999—2005年间出现了不少浅同态加密方案，例如文献[6]、[16-18]中提到的方案。目前最为著名的浅同态加密方案当属Boneh[5]等基于理想成员判定困难假设设计的加密方案。该方案能执行一次乘法和若干次加法运算，Boneh[5]等虽然用它成功解决了2DNF问题，但是该方案在解密时需要搜索解密，因此基于此方案的2DNF保密计算协议效率很低。

虽然此类加密系统为实现全同态加密方案的设计奠定了一定的基础，但是只能用于解决某些专门的问题，即能够解决的应用问题有限，很难将其拓展并且应用于解决更广泛的问题。

2.3 全同态加密方案

2009年Gentry[7]设计了首个全同态加密方案，这一里程碑事件激起了全同态研究的热潮。到目前，全同态加密方案按照构造思想大致可以分为以下3代。

（1）以Gentry[7]设计方案为代表的、基于格上困难问题构造的第1代全同态加密方案，这类方案的设计思想大致如下：

- 设计一个能够执行低次多项式运算的浅同态加密算法。
- 控制密文噪声增长，即依据稀疏子集和问题对解密电路执行“压缩”操作，然后再执行自己的解密函数实现同态解密，从而能够达到降噪的目的。
- 依据循环安全假设（即假定用方案的公钥加密自身密钥作为公钥是安全的）实现纯的全同态加密。

（2）以Brakerski-Vaikuntanathan[9]为代表的、基于带误差学习或环上带误差学习困难问题构造的第2代全同态加密方案，该类方案的构造思想大致如下：

- 归约的基础是误差学习或环上带误差学习困难问题。
- 用向量表示密钥与密文。
- 用密钥交换技术来约减密文的膨胀维数，以达到降噪目的。

该类方案的优点是不再需要电路自举技术，突破了Gentry的设计框架，在效率方面实现了很大的提升；其缺点是在使用密钥交换技术时需要增加大量用于密钥交换的矩阵，从而导致公钥长度的增长。

（3）以 Gentry-Sahai-Waters [12]为代表的、基于带误差学习或环上带误差学习困难问题构造的第 3 代全同态加密方案，此类方案的构造思想大致如下：

- 方案的安全性最终归约到带误差学习或环上带误差学习的困难问题上。
- 使用近似向量方法表示私钥，即用户的私钥实际就是密文的近似特征向量。
- 密文的同态计算使用的是矩阵的乘法与加法运算。

这类方案被认为是目前最为理想的方案，它们不再需要密钥交换与模转换技术。

3 同态加密的应用

同态加密技术在分布式计算环境下的密文数据计算方面有着广泛而重要的应用。

（1）安全云计算与委托计算。同态技术在该方面的应用可以使得我们在云环境下，充分利用云服务器的计算能力，实现对明文信息的运算，而不会有损私有数据的私密性。例如医疗机构通常拥有比较弱的数据处理能力，而需要第三方来实现数据处理分析以达到更好的医疗效果或者科研水平，这样他们就需要委托有较强数据处理能力的第三方实现数据处理（云计算中心），但是医院负有保护患者隐私的义务，不能直接将数据交给第三方。在同态加密技术的支持下，医疗机构就可以将加密后的数据发送至第三方，待第三方处理完成后便可返回给医疗结构。整个数据处理过程、数据内容对第三方是完全透明的。

（2）远程文件存储。用户可以将自己的数据加密后存储在一个不信任的远程服务器上，日后可以向远程服务器查询自己所需要的信息，远程服务器用该用户的公钥将查询结果加密，用户可以解密得到自己需要的信息，而远程服务器却对查询信息一无所知。这样做还可以实现远程用户数据容灾。

（3）密文检索。密文检索有很多方案，例如文献[7]、[19-20]中介绍的就是基于同态加密技术设计的密文检索算法。我们仅介绍 Gentry [7]用全同态加密方案实现密文检索算法：用户将要检索的内容 $f_1,f_2,\cdots,f_n$ 用公钥加密成密文 $c_1,c_2,\cdots,c_n$。将搜索引擎的查询函数置为电路 C，则搜索引擎就可以利用全同态算法中的 *Evaluate* 函数执行同态运算：

$$C_\Sigma = Evaluate(pk, C_i, (c_1, c_2, \cdots, c_n)) \quad (1)$$

其中 $C_i \in C$ 用于计算输出的第比特。当服务器将 C_Σ 返给用户时，用户用自己的私钥解密 C_Σ 即得到自己要检索的内容，而搜索引擎服务器却对用户要检索的内容一无所知。

（4）安全多方计算协议设计的工具。所谓安全多方计算就是分别持有私有数据 $x_1,x_2,\cdots,x_n$ 的 n 个人，在分布式环境中协同计算函数 $f(x_1,x_2,\cdots,x_n)$ 而不泄露各方的私有数据。以同态技术支撑的密态数据计算不仅可以满足安全多方计算协议设计中保护各方隐私的需要，还能避开不经意传输协议而大大提升协议效率。近些年来，出现了很多高效的基于同态加密的安全多方计算协议，例如文献[6]、[21]。同态加密技术已经成为安全多方计算协议设计的一个强有力的工具[22]。

（5）电子选举。基于同态加密技术设计的电子选举方案，因在计票快捷与准确、节省人力与开支、投票的易用性等方面较传统投票方式有着无法企及的优越性，越来越受到人们的青睐。目前出现了很多基于同态的电子选举方案，像文献[23-24]中介绍的都是基于同态的电子选举方案。在此描述一下 Damgård [23]方案：选民将自己的选票 $v_i \in \{0,1\}$ 加密 $c_i = Enc(v_i)$，选票中心统计部门（拥有密钥且可信）利用同态加密方案的同态性统计计算选票数 $V = Dec(c_1 \times c_2 \times \cdots c_n) = v_1 + v_2 + \cdots + v_n$。

（6）其他方面的应用。同态加密技术在其他方面也有诸多应用，例如多方零知识证明、软件保护、聚合（同态）签名等。

4 同态技术存在的问题

综上所述，目前同态技术及其应用方面还存在以下几个重点问题需要我们进一步研究。

（1）只能实现单比特加密，如何高效地实现全同态加密有待进一步研究。

（2）大多基于未论证的困难问题，寻找可论证的困难问题依然是个摆在密码工作者面前的难题。

（3）大多只能达到 IND-CPA 安全，偶有能达到 IND-CCA1 安全，但还未见能达到 IND-CCA2 安全的，同态加密系统的安全性研究有待进一步提高。

（4）需要额外的消除噪音算法，依然不是自然同态，如何设计一个具有自然同态性的全同态加密方案依然是一个开问题。

（5）在安全云计算、安全多方计算、密态数据计算等领域的应用还处在初级阶段，有待于进一步地拓展；同时，在其他领域的相关应用也需要积极开拓。

5 结束语

随着分布式计算的普及，如何保护分布式计算环境下的用户隐私已经成为一个重要问题，密码学中的同态加密技术可以为分布式计算环境的用户隐私保护提供强有力的技术支撑。文章介绍了同态加密技术的研究现状及研究展望，并简单介绍了基于同态加密技术的密态数据计算在分布式计算中的各种应用。目前，同态加密方案在安全性方面大都只能达到 IND-CPA 安全，如何设计更高安全级别的同态加密方案依然需要进一步研究。全同态加密方案的构造还处于理论阶段，尚不能用于实际

的密态数据计算问题，如何设计基于代数系统的（自然）全同态加密方案依然是未来研究的重点。同态加密技术支撑的密态数据计算在其他领域的应用有待进一步拓展。

参考文献

[1] RIVEST R L, ADLEMAN L, DDRTOUZOS M L. On Data Banks and Privacy Homomorphisms [J]. Foundations of secure computation, 1978, 4(11): 169–180
[2] RIVEST R L, SHAMIR A, ADLEMAN L. A Method for Obtaining Digital Signatures and Public–Key Cryptosystems [J]. Communications of the ACM, 1978, 21(2): 120–126. doi: 10.1145/359340.359342
[3] ELGAMAL T. A Public Key Cryptosystem and a Signature Scheme Based on Discrete Logarithms [M]. Advances in cryptology. Germany: Springer Berlin Heidelberg, 1985: 10–18. doi: 10.1007/3–540–39568–7_2
[4] OKAMOTO T, UCHIYAMA S. A New Public–Key Cryptosystem as Secure as Factoring [M]. Advances in Cryptology—EUROCRYPT' 98. Germany: Springer Berlin Heidelberg, 1998: 308–318. doi: 10.1007/BFb0054135
[5] PAILLER P. Public–Key Cryptosystems Based on Composite Degree Residuosity Classes [M]. Advances in Cryptology—EUROCRYPT'99. Germany: Springer Berlin Heidelberg, 1999: 223–238. doi: 10.1007/3–540–48910–X_16
[6] BONEH D, GOH E J, NISSIM K. Evaluating 2–DNF Formulas on Ciphertexts [M]. Theory of Cryptography. Germany: Springer Berlin Heidelberg, 2005: 325–341.doi: 10.1007/978–3–540–30576–7_18
[7] GENTRY C. Fully Homomorphic Encryption Using Ideal Lattices[C]//Proceedings of the Forty–First Annual ACM Symposium on Theory of Computing, 2009: 169–178. doi: 10.1145/1536414.1536440
[8] VAN D M, GENTRY C, HALEVI S, et al. Fully Homomorphic Encryption Over the Integers [M].Advances in Cryptology–EUROCRYPT 2010. Germany: Springer Berlin Heidelberg, 2010: 24–43
[9] BRAKERSKI Z, VALIKUNTANATHAN V. Fully Homomorphic Encryption from Ring–LWE and Security for Key Dependent Messages [M]. Advances in Cryptology–CRYPTO 2011. Germany: Springer Berlin Heidelberg, 2011: 505–524
[10] BRAKERSKI Z, GENTRY C, VAIKUNTANATHAN V. (Leveled) Fully Homomorphic Encryption without Bootstrapping [C]//Proceedings of the 3rd Innovations in Theoretical Computer Science Conference, 2012: 309–325.doi: 10.1145/2090236.2090262
[11] BRAKERSKI Z. Fully Homomorphic Encryption without Modulus Switching from Classical GapSVP [M]. Advances in Cryptology–CRYPTO 2012. Germany: Springer Berlin Heidelberg, 2012: 868–886. doi: 10.1007/978–3–642–32009–5_50
[12] GARG S, GENTRY C, HALEVI S, et al. Attribute–Based Encryption for Circuits from Multilinear Maps[M]. Advances in Cryptology–CRYPTO 2013. Germany: Springer Berlin Heidelberg, 2013: 479–499. doi: 10.1007/978–3–642–40084–1_27
[13] BRAKERSKI Z, VALIKUNTANATHAN V. Efficient Fully Homomorphic Encryption from (standard) LWE [J]. SIAM Journal on Computing, 2014, 43(2): 831–871
[14] CHEON J H, KIM J, LEE M S, et al. CRT–Based Fully Homomorphic Encryption over the Integers[J]. Information Sciences, 2015, 310: 149–162
[15] GOLDWASSER S, MICALI S. Probabilistic Encryption [J]. Journal of computer and system sciences, 1984, 28(2): 270–299
[16] I FERRER J D. A New Privacy Homomorphism and Applications[J]. Information Processing Letters, 1996, 60(5): 277–282. doi: 10.1016/S0020–0190(96) 00170–6
[17] DOMINGO–FERRER J. A Provably Secure Additive and Multiplicative Privacy Homomorphism*[M]. Information security. Germany: Springer Berlin Heidelberg, 2002: 471–483
[18] MELCHOR C A, GABORIT P, HERRANZ J. Additively Homomorphic Encryption with D–Operand Multiplications [M]. Advances in Cryptology–CRYPTO 2010. Germany: Springer Berlin Heidelberg, 2010: 138–154. doi: 10.1007/978–3–642–14623–7_8
[19] LI J, WANG Q, WANG C, et al. Fuzzy Keyword Search over Encrypted Data in Cloud Computing[C]//INFOCOM, 2010 Proceedings IEEE, 2010: 1–5
[20] HU H, XU J, REN C, et al. Processing Private Queries Over Untrusted Data Cloud Through Privacy Homomorphism[C]// 2011 IEEE 27th International Conference on Data Engineering (ICDE), 2011: 601–612
[21] 李顺东, 王道顺. 基于同态加密的高效多方保密计算[J]. 电子学报, 2013, 41(4): 798–803
[22] BENDLIN R, DAMGARD I, ORLANDI C, et al. Semi–Homomorphic Encryption and Multiparty Computation[M]. Advances in Cryptology–EUROCRYPT 2011. Germany: Springer Berlin Heidelberg, 2011: 169–188
[23] CRAMER R, GENNARO R, SCHOENMAKERS B. A Secure and Optimally Efficient Multi–Authority Election Scheme [J]. European Transactions on Telecommunications, 1997, 8(5): 481–490
[24] DAMGARD I, JURIK M. A Generalisation, a Simplication and Some Applications of Paillier's Probabilistic Public–Key System [C]//Public Key Cryptography. Springer Berlin Heidelberg, 2001: 119–136.doi: 10.1007/3–540–44586–2_9

作者简介

巩林明，陕西师范大学在读博士研究生；主要研究方向为信息安全与密码学；发表SCI、EI检索论文2篇。

李顺东，陕西师范大学教授、博士生导师；先后主持国家“863”高技术发展项目、国家自然基金项目和国际合作项目多项；发表SCI、EI检索论文30余篇。

郭奕旻，陕西师范大学在读硕士研究生；主要研究方向为信息安全与密码学；发表SCI、EI检索论文3篇。

综合信息

全球光纤连接器市场2020年复合年增长率将达9.9%

据专业机构预测：全球光纤连接器市场预计将在2020年达到49亿美元，2015—2020年的复合年增长率将达到9.9%。

这种增长可以归因于更高带宽的应用，需要使用光纤电缆和连接器来保证带宽的安全性和高速。显然，电信和数据应用，例如云端、音频、视频、电视和在线游戏都是光纤市场的巨大推动力。同时航空航天和国防等领域的工业应用，对光纤连接器来说是一个更大的市场。

光纤连接器在安全系统中的应用，预计在2015—2020年间的复合年增长率将达到12.4%，预计将为这一市场的参与者提供潜在的增长机会。

（转载自《中国信息产业网》）

DOI:10.3969/j.issn.1009-6868.2016.01.008
网络出版地址:http://www.cnki.net/kcms/detail/34.1228.TN.20151117.1459.002.html

计算机网络取证和调查的科学研究

Computer Network Forensics and Investigation

邹锦沛/KP CHOW[1]
陈航/CHEN Hang[2]
徐菲/XU Fei[3]

(1.香港大学 计算机科学系,香港 999077;
2. 南京理工大学 计算机学院,江苏 南京,210094;
3. 中国人民大学 法学院,北京 100872)
(1. Department of Computer Sciences, Hong Kong University, Hong Kong 99077, China;
2. Department of Computer Sciences, Nanjing University of Science and Technology, Nanjing 210094, China;
3. Law School, Renmin University of China, Beijing 100872, China)

中图分类号:TN929.5 文献标志码:A 文章编号:1009-6868 (2016) 01-0030-004

摘要: 认为针对计算机犯罪,现代的调查是对电子证据进行智能相关性分析,并发掘同一事件不同证据之间的联系;而证据分析又包括电子数据证据的分析、对收集的数据和备份进行查找、分析、归类,以及犯罪现场重建等。提出犯罪现场重建是计算机网络犯罪调查的重要部分。通过理论和实验分析,将取证科学应用到网络犯罪调查上,并以P2P网络调查作为例子,分析如何通过调查取证来寻找数据的第一个上传者。认为只有将恰当的法证科学适时应用到电子证据取证调查中,才能够更好地重构犯罪场景,还原案件真相并实现法律正义。

关键词: 电子证据;调查;法证科学;犯罪现场重构

Abstract: For computer related crimes, modern digital investigation emphasizes analysis of the relationship between different digital evidence with the goal of determining how different pieces of digital evidence appear in a single event. Digital evidence analysis includes searching, classification, analysis, and event reconstruction. Event reconstruction is the most important part of cybercrime investigation. By theoretical and experimental analysis, one can apply forensic science to cybercrime scene reconstruction. We demonstrate how to apply forensic science in cybercrime investigation involving the peer-to-peer network, with the objective of identifying the first uploader in the peer-to-peer network. By applying forensic science to cybercrime investigation, the digital investigator should be able to reconstruct the crime scene more efficiently.

Key words: digital evidence; investigation; forensic science; crime scene reconstruction

1 计算机法证的概况

1.1 电子证据与取证调查

人们常把由计算机制作的文件和计算机活动日志当作电子证据。根据香港特别行政区的《证据条例》第22A条[1],一项由计算机制作的文件的陈述,由在有关计算机的运作或有关活动的管理方面身居要职的人,依法签署的证明书依法证明后,则可在任何刑事法律程序中,接纳为该陈述内所述任何事实的表面证据。该证明书对由计算机制作的文件的制作方式予以描述,并在有关法律程序关系的范围内,说明该文件的性质及内容。

现代的计算机网络罪行更为复杂[2-4]。犯罪分子不仅窜改计算机记录,还用计算机来存储他们的数据,例如:非法金融交易和地址簿。此外,犯罪分子还利用互联网作为犯罪平台,例如分布式拒绝服务攻击、网络拍卖诈欺、分享受版权保护的作品等。现代计算机网络犯罪的主要特点是专业性强、有组织、并利用网络。

传统的计算机法证取证技术,已无法应对当今的计算机网络犯罪。现代的调查是对电子证据进行智能相关性分析,并发掘同一事件不同证据之间的联系。现代的分析证据是指对电子数据证据的分析,对收集的数据和备份进行查找、分析、归类等。

1.2 计算机法证的发展

计算机法证成立于20世纪70年代,其发展阶段可分为:婴儿期、儿童期和青春期。婴儿期为1985—1995年,儿童期为1995—2005年,青春期为2005—2010年[6-9]。计算机法证的研究重点是资料恢复,其中最主要是数据恢复、密码恢复和文件恢复[10]技术。数据恢复指恢复已被移走或删除的电子逻辑或物理数据,例如一个破碎的硬盘;密码恢复指处理受密码保护的原始数据,如密码加密的文件;文件恢复则尝试从硬盘内的数据块的片段恢复删除的文件。目前,文件恢复技术在计算机法证工作中,是一个主要的工作内容,例如手机的数

收稿日期:2015-11-08
网络出版日期:2015-11-17

据恢复[11]。

传统的法证主要集中在识别和重建。识别包括指纹、DNA 和毒品。指纹和 DNA 被用来识别特定的人，而毒品分析用于确定毒品的化学成分。识别的目的是用来判断样品是否来自一个特定的对象，诸如人或毒品。重建[12]则包括犯罪现场重建与弹道重建。犯罪现场重建指试图重建在犯罪现场所发生的事件，例如重现一宗谋杀案如何发生；而弹道的重建被用来重建从枪炮发射的子弹的轨迹。

计算机法证与传统法证相似，它们都试图解答一些问题，如：发生了什么事情？当事人是谁？什么时间、什么地点、如何发生的？这件事情发生的动机是什么？

2 计算机法证犯罪现场重建

在香港特别行政区及世界各地，藏有儿童色情物品是一种犯罪行为。香港地区某犯罪嫌疑人被指控藏有儿童色情物品，警方查获计算机一台。

2.1 犯罪现场重建过程

为了进行计算机法证分析，警方按标准的程序检查检获的计算机。这些标准程序基于确立的计算机采证程序，并产生法证克隆、计算哈希算法值、扫描计算机病毒等等。在标准的计算机采证过程后，计算机法证鉴定人将会分析检获的计算机硬盘驱动器，并收集电子证据。对于儿童色情物品的案件，电子证据便是儿童色情图片和动态影像。所以第一步是寻找在计算机内储存的儿童色情图片和动态影像，这些会包括：现存的儿童色情图片和动态影像、恢复删除的儿童色情图片和动态影像。图 1 显示了恢复删除的文件。

执法人员已为儿童色情图片和动态影像建立哈希值数据库，计算机法证鉴定人便不需要阅览个别的儿童色情图片和动态影像。相反，计算机法证鉴定人只需要把计算机里面的文件的哈希值与数据库比较。图 2 显示出根据该文件的创建时间和最后写入的时间的事件恢复过程。根据这两个时间戳，文件在 2005 年 2 月 22 日上午 12 时 37 分 09 秒被复制到当前位置。

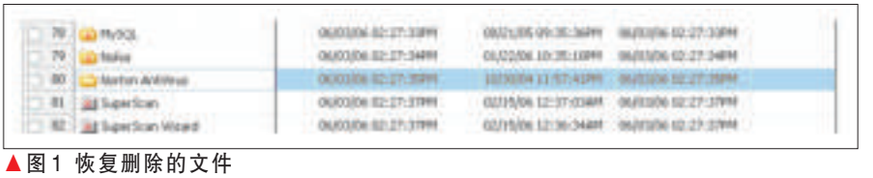
计算机法证鉴定人试图从文件中包含的关于文档的信息，例如创建者、修改的日期和其他细节，和其他计算机活动日志和电子证据，去重建产生该儿童色情图片和动态影像的经过。例如 2004 年 1 月 4 日，疑犯从互联网下载童色情图片和动态影像，并使用信用卡号码 0000-1111-2222-3333 在网上支付，然后在 2009 年 8 月 5 日备份到外部媒体。

图 3 显示了犯罪现场重建，犯罪嫌疑人在 2004 年 1 月 2 日使用他的信用卡购买在互联网上的儿童色情物品，然后在 2009 年 8 月 4 日做一个备份到外部硬盘上。

2.2 犯罪现场重建的法证科学

犯罪现场重建之所以重要，是因为在法庭上，计算机法证鉴定人需要给非技术人员（法官和陪审团）解释“技术细节”，让他们做出裁决。法律可能有特定的要求，例如控方需要证明犯罪嫌疑人知悉儿童色情图片和动态影像储存在计算机里。许多时候，犯罪嫌疑人的辩解是，儿童色情裸照是被木马下载的，他并不知道它们为什么在计算机里。辩方可以聘请专家质疑检察官的说法，或混淆法官和陪审团对“技术细节”的了解。计算机法证鉴定人的证言不仅受到辩方的质疑，还受到法庭的监察。

美国最高法院在 Daubert 一案裁定，主审法官必须确保接纳的科学证言或证据，不但全部切题并且全部可靠。所提出的科学证言须以恰当的



▲图 1 恢复删除的文件

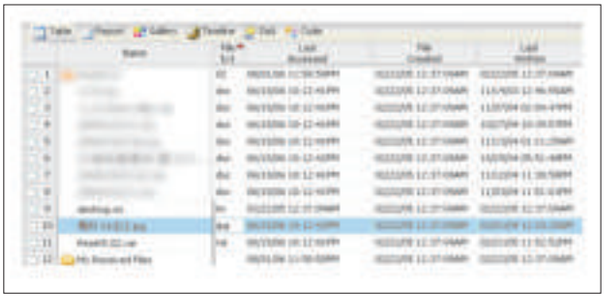


图 2▶ 事件重建

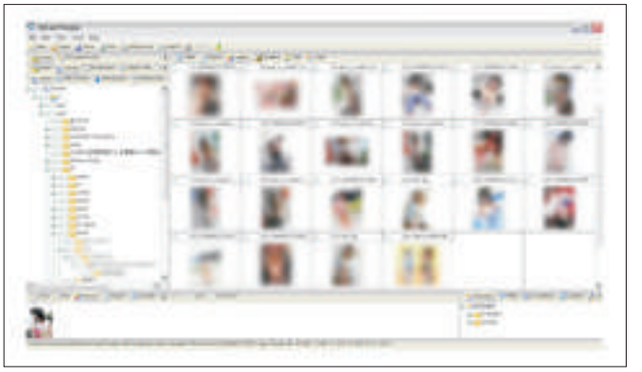


图 3▶ 犯罪现场重建

验证方法，获科学上已知的有力理据支持，具体包括：

（1）通过可靠性测试。

（2）通过同行评审。

（3）提供方法或理论的错误率，并在一定范围之内。

（4）符合标准和控制。

（5）得到普遍接受。

有关的法律也规定专家的科学证言须结合“科学知识”，并就证据的可靠性及可信性立下标准。证据的可靠程度取决于科学上能否验证。

现代的电子调查是对电子证据进行智能相关性分析并重建犯罪过程，例如通过计算机的所有者、电子签名、密码、交易纪录、邮件、发送服务器的日志、上网IP等计算机特有信息识别体，同其他证据互相印证、相互关联，然后进行综合分析。

同时，很多时候电子证据还需要传统的调查手法相辅助。调查人员在重建罪行时，需要适当考虑其他可能存在的解释，并在解释与证据之间进行相互印证。可能在某种假设情况下，需要查找更多的证据；又可能在新的证据下，得出新的假设和解释。调查人员要把证据互相印证，相互关联起来进行综合分析。调查人员要找出与理论假设与证据之间如何构成证实的关系，才能准确地重构犯罪过程。构建假设并验证就是可以采用的科学方法。

3 P2P网络中发布者取证调查

我们以Foxy软件为例，介绍在一个点对点分享（P2P）网络中，如何通过调查取证找到数据的上传者，以及何时调查能够找到数据的最先上传者。2008年香港艳照门事件中，嫌疑人就是利用Foxy对艳照进行共享，使得艺人的裸照在Foxy网络中迅速传播。想要抓捕嫌疑人，需要通过对Foxy进行分析，找到物理世界中的人。

Foxy是一个繁体中文P2P软件，发行者为一家已于2010年关闭的台湾公司。该软件只有正体中文版，没有英文等其他语言的版本，因此主要流行在台湾、香港及澳门等使用繁体中文的地区。它利用强制上传增加分享速度，但用户无法停止上传。它没有路由机制，源头的私隐（例如IP位址，所在地点）不受保障。它没有连接加密，连接容易被监视。它容易让使用者误设为全机分享，分享用户所有档案，每次下载完成后会自动重新分享用户的所有档案，并且用户无法停止。

当Foxy客户端试图连接到Foxy网络，会执行以下任务，如图4所示。

（1）用户连接到Foxy的服务器，以获得一个对等端列表。

（2）服务器回答一个对等的用户列表。

（3）用户发送一个PING请求到各个对端。

（4）各个对等端回答一个PING请求到用户。

（5）用户现在在Foxy网络上。

在Foxy网络中，每个共享文件都使用它的名字。当一个用户要寻找一个文件，他输入了一个搜索查询的文件名（或只是其中的一部分）。查询信息发送到所有对等端，然后传递给其他相邻对等端。当一个对等端拥有一个文件名和查询信息字符串相匹配，就回答一个QueryHit消息给发出请求的用户。QueryHit消息中包含的信息包括IP地址和端口号，共享文件和文件本身的信息。这使用户能够建立一个连接到该对等端，并启动下载。

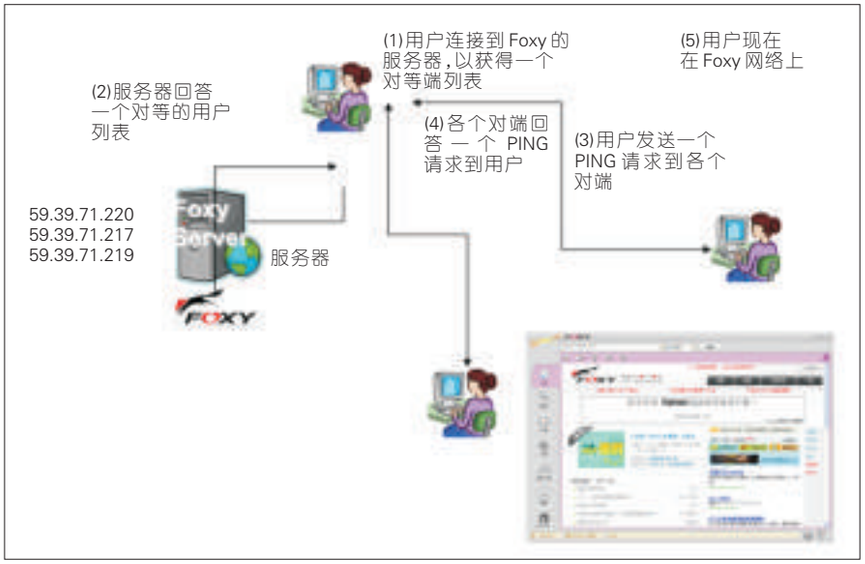
在接收QueryHit信息以及有关的连接信息，用户需要先选择一个文件，并发送一个TCP HTTP GET至承载该文件的等端，请求下载。承载该文件的等端然后回应，并开始发送所请求的数据。

与所有对等网络的文件共享，所有等端在Foxy网络中都是同等级的。所有拥有与“Query”请求匹配的副本的等端都回复它的IP地址给请求者。图5显示了在P2P网络中文件的分发种子数据增长曲线。当一个文件被广泛地分布后想要确认哪个等端是发起者多是不可能的。也许可以在下列情形中找到：

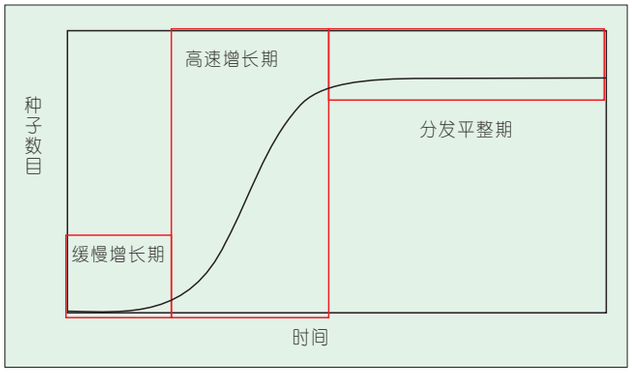
（1）在peer缓慢的增长期。

（2）分享文件非常大。

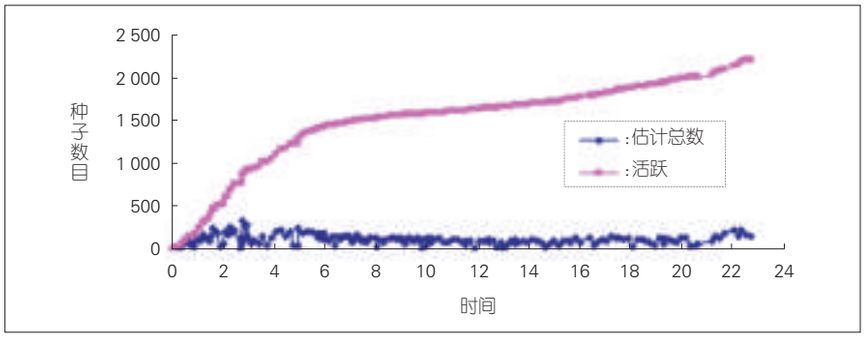
在缓慢增长期后找到谁是发起者也是不可能的。在连续监察下，我们认为如果识别到大量关于特定名字的查询，并且发现大量查询命中来自同一个IP地址，则很有可能该IP



▲图4 Foxy客户端连接到Foxy网络



◀图5
在P2P网络中的文件分发



▲图6 实验确认缓慢上升期的存在

地址就是第1个上传者。我们进行了一系列相关的实验，实验结果验证了缓慢增长期的存在。实验结果如图6所示。

对于网络上数据的上传者，需要进行如下的科学分析：

（1）查清网络物品的来源。

（2）查清“虚拟嫌疑人”。通过上传人IP地址（包括静态IP、动态IP）、上网账号、密码及其相关登记资料、通过关联点分析网上活动轨迹及对资讯内容的分析，判断行为人的网络行为、个性特征，锁定虚拟嫌疑人。

（3）确认“现实嫌疑人”。这一般属于通常所说的落地调查，即通过询问嫌疑人，并通过现场搜查、勘验、检查及电子数据鉴定确定现实嫌疑人，其中硬盘、手机、日志、光盘的电子数据的固定和提取尤为重要。

4 结束语

现代计算机网络犯罪调查的重要部分是犯罪现场重建。在计算机网络取证中，电子证据在犯罪现场重建中往往能够起到关键作用。运用各种科学手段、方法和技术，分析电子证据中隐含的信息及线索，能够更好地重新构建或模拟一个犯罪现场。我们相信计算机网络犯罪调查既是一门技术，更是一门科学，只有将恰当的法证科学适时应用到电子证据取证调查中，才能够更好地重构犯罪场景，还原案件真相并实现法律正义。

参考文献

[1] KWAN M, OVERILL R E, CHOW K P, et al. Sensitivity Analysis of Digital Forensic Reasoning in Bayesian Network Models [C]// Proceeding of 7th Annual IFIP WG 11.9 International Conference on Digital Forensics, Orlano, Florida, USA, 2011
[2] 钟琳, 黎家盈, 邹锦沛, 等. 基于多视图分析的复杂网络犯罪现场重构[J]. 电信科学, 2010, (S2):165–170
[3] LAW F, LAI P, CHOW K P, et al. Memory Acquisition: A 2–Take Approach [C]//The 2009 International Workshop on Forensics for Future Generation Communication environments (F2GC–09), Jeju Island, Korea, Dec 10 – 12, 2009
[4] FRANK Y W, LAW, CHOW K P, et al. A Host–Based Approach to BotNet Investigation [C]// Proceeding of the 1st International Conference on Digital Forensics and Cyber Crime, Albany, NY, Sept 30–Oct 2, 2009
[5] HE Y, ZHANG P, HUI C K, et al. Cloud Forensics Investigation: Tracing Infringing Sharing of Copyrighted Files in Cloud [C]// Proceeding of 2012 ADFSL Conference on Digital Forensics, Security and Law (ADFSL’12), 30–31 May 2012
[6] XU F, CHOW K P, He J, et al. Privacy Reference Monitor–A Computer Model for Law Compliant Privacy Protection [C]//2009 IEEE International Conference on Parallel and Distributed Systems, Shenzhen, China, 2009
[7] PUN K H, HUI L C K, CHOW K P, et al. Review of the Electronic Transaction Ordiance, Can the Personal Identification Number Replace the Digital Signature [J]. Hong Kong Law Journal, 2002, 32(2):241–257
[8] IEONG S C R, CHOW K P. Enhanced Monitoring Rule Through Direct Node Query for Foxy Network Investigation[C]// The First International Conference on Digital Forensics and Investigation (ICDFI), Beijing, China, 2012
[9] YE Y, WU Q, LI Y, CHOW K P, et al. Unknown Chinese Word Extraction Based on Variety of Overlapping Strings [J]. Information Processing and Management, 2013, 49(2) : 497–512
[10] CASEY E. Digitla Evidence and Computer Crime: Forensic Science, Computers and the Internet [J].Jurimetrics, 2011, 11(3): 373
[11] 戴士剑. 电子证据调查指南[M]. 北京: 中国检察出版社, 2014
[12] LAI P K Y, CHOW K P, HUI L, et al. Modelling the Initial Stage of a File Sharing Process on a BitTorrent Network [J]. Peer–to–Peer Networking and Applications, 2014, 7(4): 311–319

作者简介

邹锦沛，香港大学副教授；主要研究领域为网络安全与电子取证；先后主持和参加项目20余项，获得多项科研成果奖；出版专著、已发表国际会议、期刊论文50余篇，其中被SCI/EI检索40余篇。

陈航，南京理工大学硕士；主要研究领域为内部威胁分析与电子取证技术；参加项目2项，获得了多项研究成果。

徐菲，中国人民大学博士后；主要研究领域为网络安全、隐私保护与电子取证；先后主持和参加项目10余项；已申请发明专利、软件著作权10余项，发表国际会议、期刊论文30余篇，其中被SCI/EI检索20余篇。

DOI:10.3969/j.issn.1009-6868.2016.01.009
网络出版地址:http://www.cnki.net/kcms/detail/34.1228.TN.20151216.2117.002.html

动态网络主动安全防御的若干思考

Proactive Security Defense of Dynamic Network

中图分类号:TN929.5 文献标志码:A 文章编号:1009-6868 (2016) 01-0034-004

摘要: 提出以动态化、随机化、主动化为特点的动态网络主动安全防御是解决信息系统中未知漏洞与后门攻击的一种新途径。在动态网络的主动变迁技术中,提出了演进防御机制(EDM),该机制可以根据网络系统安全状态、网络系统安全需求等,选择最佳的网络配置变化元素组合来应对潜在的攻击、保证特定等级的安全要求。网络的动态重构和变迁需要根据系统的安全态势和可能遭受的网络攻击来考虑,其关键是如何有效对系统的安全态势和网络的攻击进行主动探测与感知。尚处于起步阶段的动态网络主动安全防御的创新技术研究任重而道远。

关键词: 被动防御;未知攻击;主动防御

Abstract: Dynamic network proactive security defence is an effective method for solving the unknown vulnerabilities and back door attacks in the information system. In this paper, the evolution defense mechanism (EDM) is proposed. According to the security status, network system security needs, EDM can select the best network configuration change elements to deal with potential attacks and ensure the safety requirements of a specific level. Dynamic reconfiguration and changes of the network need to be considered in accordance with the security situation of the system and the possible network attacks. The key is how to detect and the security situation and network attacks. It proposes that study on dynamic network proactive security defense in China is in the initial stage, and there is still much work to do.

Key words: passive defense; unknown attack; proactive defense

吴春明/WU Chunming
(浙江大学 计算机系统结构与网络安全研究所,浙江 杭州 310027)
(Institute of Computer System and Network Security, Zhejiang University, Hangzhou 310027, China)

- 动态网络主动安全防御技术已成为信息安全领域的重要研究方向
- 大数据中心、云平台等技术设施的出现,能使我们更有效地进行系统的数据采集、处理和分析
- 主动防御机制是对目前被动防御系统的拓展和补充,两者能够相辅相成

棱镜计划的曝光和斯诺登事件的持续发酵,表明了美国网络监控计划和针对中国物理隔离网络的威胁已穿透国门直逼我们面前,以“物理隔离”为基础的最后一道防线不复存在。

传统的网络安全防御思想是在现有网络体系架构的基础上建立包括防火墙和安全网关、入侵检测、病毒查杀、访问控制、数据加密等多层次的防御体系来提升网络及其应用的安全性。但近年来不断被披露的网络安全事件及由此带来的严重后果也逐渐暴露了传统的网络安全防御技术存在的问题,尤其是难以有效抵御系统未知软硬件漏洞的攻击,难以防御潜在的各类后门攻击,难以有效应对各类越来越复杂和智能化的渗透式网络入侵。近年来,研究动态网络主动安全防御技术已成为信息安全领域的重要方向。

如何通过动态化、随机化、主动化的手段改变网络信息系统的运行或执行环境,突破传统网络信息安全被动防御的窘境,将“亡羊补牢”式的被动防御转变为难以被侦测的主动防御是一种值得我们思考的创新思路[1-3]。

中国作为信息技术后进国家,关键领域的信息基础设施或装备几乎都被西方寡头公司控制,绝大部分核心元器件、高端芯片、基础/工具软件和半导体制程装备也都严重依赖进口。中国网络空间几乎单向透明地呈现在以美国为首的西方发达国家面前,信息安全面临极大威胁。当前,网络空间的软件攻击技术发展迅速,且已进入自动化程度高、攻击速度快、攻击工具复杂、潜伏更趋隐蔽的高级阶段。集成电路设计、制造、封装、测试和工具软件等环节被全球少数寡头企业垄断,在硬件或物理电路中植入或预留后门/陷阱成为新的攻击手段,且难以检测和有效防范。正在兴起的软硬件和电磁协同攻击技术具有植入更隐蔽、识别更困难、

收稿日期:2015-11-20
网络出版时间:2015-12-28

隔离效果差、清理代价高等综合优势，正在成为主流攻击模式，给安全风险的评估、检测和防御都带来了前所未有的挑战，甚至作为防御方底线的加密手段也很难保证其最终的有效性，现有网络空间安全手段的匮乏已严重危及国家主权和社会安定。

在全球经济一体化、专业分工国际化的时代背景下，任何一个国家都不可能完全掌控包含设计链、生产链和供应链在内的完整的安全链。从理论上讲，目前软硬件设计中的漏洞又是不可避免的，即使中国自主设计的芯片、硬件和软件，迄今尚无有效的漏洞检测手段和杜绝方法，从工程科学的角度也很难证明在整个安全链诸环节中没有漏洞、或未被植入病毒木马和预留后门。而近年来，控制系统安全事故频发，中国工业安全也受到了严重威胁，已引起了各国政府和学术界的广泛关注[4]。

动态网络主动安全防御以提供运行环境的动态性、非确定性、异构性、非持续性为目的，通过网络系统中的环境、软件、数据等主动重构或迁移实现动态环境，以防御者可控的方式进行主动变化，对攻击者则表现为难以观察和预测的动态目标，从而大幅增加攻击难度和成本，大幅降低系统安全风险[5]。

文章从多个方面论述了我们对动态网络主动安全防御技术的一些思考和目前所做工作。

1 动态网络的主动变迁技术

安全这个话题，多年来是从防护、抵御外界的入侵、攻击的角度来谈论的，但由于目前防御和攻击之间的不对称，因此一旦遭到无所不在的攻击后，能主动招架、有效抵抗的可能性就变得很小。据此形态，我们可以从另外一个角度来看待这个问题：即利用网络的基础设施、传输协议、数据访问等能力来完成特定的某项任务，在执行任务过程中有效保护好各环节的动作或操作，对注入、驻留、渗透的攻击所需感知或匹配的网络环境进行动态变换，以切断攻击行为过程中的先验知识链（即在平时的攻防僵持状态，双方保持静态化）；因没有实质性的行为发生，此时的威胁是可容忍的；在访问、请求等功能或操作启动后即可发起动态变换的指令，配置成与静态化对峙时刻不同的网络状态，增加入侵攻击的难度；一旦任务结束即可释放各类资源，形成机动网络的生成、配置、执行与释放卸载的动态化机制[6]。

为了增强网络动态变化的随机性，可以增加部分冗余网络节点设备和链路，在不同时刻根据不同需求将其随机激活或休眠，以动态重构网络，从而使得整个网络难以被探测。

通过网络与配置的动态化，将静态的网络变成动态、随机可重构的网络，改变目前网络攻防双方不对称的状况，使攻击者无法有效确定攻击目标，从而实现动态的网络安全。

动态网络的主动重构防御的指导思想是：在保障网络服务、功能等价的前提下，利用网络可重构的技术手段，构建具有依据任务需求、主动变迁网络运行与传输环境的网络架构，通过对破坏网络攻击链的方法及措施的研究和设计，提高网络攻击难度，形成主动网络防御的能力。

动态网络的主动变迁需要在考虑现有网络基础设施组成结构的基础上，结合拓扑结构、路由、环境、软件等网络要素来考虑，其关键技术、变迁策略和协同机制可关注如下几个方面：

（1）潜伏部分网络物理节点设备。通过潜伏部分网络设备，在必要时将其激活，重构出一个网络，在不同时刻或者根据不同需求，可以激活不同的潜伏设备或休眠部分设备，从而增强探测整个网络结构的难度。

（2）按需动态生成逻辑服务网络。逻辑服务网络是按需动态生成和释放的，即使对同一用户提出的需求，在不同时段所生成的逻辑服务网络拓扑结构、映射到的物理设备位置、逻辑设备地址等都有可能不一样。因此，多样化变换的逻辑服务网络使得网络结构也更难以被探测，数据难以被追踪[7]。此外，同一物理设备内各逻辑设备间的隔离化处理，也可以有效降低诸如路由器分布式拒绝服务（DDos）攻击的危害。

（3）多样化、差异化的网络服务。通过构建逻辑服务网的方式为不同业务提供不同等级的安全传输服务，高效地重构、利用网络资源。

（4）多径、动态路由。通过动态路由，用户信息每次传输路径都可能不同，使得难以被追踪定位[8]；通过多径路由，用户信息在不同的路径上传输，即使部分信息被截获，窃听者也很难获得完整的信息。另外，通过域间/内虚拟路由技术、域间/内路由代理技术等也不失为可进一步拓展的手段。

（5）网络设备的“白盒”设计。将网络设备封闭结构下不同部件的松耦合，比如软件定义网络（SDN）实现的数据面、控制面和应用软件的分离，而可重构网络则基于开放标准的构件化设计思想，实现粒度可伸缩的弹性模块化，从而提高设备主动避免干扰的能力。

（6）IP地址的可变。通过IP地址变换，可达到资源在IP层面的动态变化，使网络扫描攻击环节失效，从而形成对后续网络攻击失去有效目标的有利局面，减缓网络中蠕虫、病毒和木马的转播[9]。

（7）SDN的控制器联动。可以考虑新型网络操作系统，采用多控制器联动、协同与虚拟配置，重构虚拟网络与虚拟机等的方式进行协同化变迁控制单元与组件。

在前期的研究工作中我们将生物启发方法[10]用于主动网络安全机制中，并且提出了演进防御机制（EDM）[11]。该机制根据网络系统安全状态、网络系统安全需求、用户特定应用的安全需求，选择最佳的网络

配置变化元素组合来应对潜在的攻击、保证特定等级的安全要求。通过结合目前SDN最新控制器技术，EDM架构的愿景可设为：保证多种动态网络配置变化元素种类的共存，避免在配置变化过程中的冲突，并充分利用SDN所具有的良好的可编程性，通过不断更新动态网络配置元素种类、更趋有效的网络配置动变策略来应对新的威胁，从而保证EDM的不断持续演进，提高处理威胁的效果、增强安全增益。同时，EDM架构的设计考虑了动态网络配置所带来的网络效能损失问题，使其能够根据网络实际特点来选择适合的动态网络配置组合策略，因而具有自适应网络环境的特性。在原理验证用例中同时实现了IP地址与数据流路径随机化变化机制，并实现了两种随机化机制的协同与冲突避免。

2 动态网络安全态势和网络攻击的感知

动态网络主动安全防御能有效防御未知漏洞和后门攻击，实现动态的网络安全，但需要考虑主动防御的代价。过于频繁的动态变化网络元素，则从安全代价权衡而言，可能是昂贵的。因此，网络的动态重构和变迁需要根据系统的安全态势和可能遭受的网络攻击来考虑，其关键是如何有效对系统的安全态势和网络的攻击进行主动探测与感知。

（1）多维数据和全局性能指标综合下的安全态势感知

可以通过在网络的各个关键节点处（包括安全防护设备、流量采集设备、关键资源服务器、交换路由设备等）部署传感器或通过数据采集接口，实时采集各节点的运行数据，包括系统告警、资源占用等，将其发送到系统的全局监控中心，进行实时分析、感知系统的当前安全态势[12]。

目前，大数据中心、云平台等技术设施的出现，一方面提高了系统的计算能力和存储能力，另一方面也能使我们可更有效地进行系统的自动化数据采集。这一点在目前的SDN和软件定义安全（SDS）[13]系统中更为突出。在SDN和SDS系统中，管控系统拥有全局视图和知识库，开放和标准化的安全接口，可协同安全设备联动，可对网络流量、网络行为、安全事件等进行自动化、全面的采集、分析。如果能通过大数据多维分析，从全局角度对威胁进行有效分析建模，则系统可根据所建立的模型有效感知系统安全态势，从而进行针对性的主动防御，增加入侵攻击的难度，提高系统的安全性[14]。

（2）利用伪装、诱骗等手段进行搅局

利用伪装、诱骗等手段，使入侵者无法得到真实的系统信息等先验知识，诱骗入侵者攻击一些预先设置的陷阱蜜罐系统，来发现入侵[15]。

- 指纹伪装和隐藏。将真实系统的指纹信息按照定制策略进行改变，返回“真真假假，虚虚实实”的系统信息，使入侵者无法确定系统的真实版本等信息，无从下手。并进行指纹伪装，返回一些不存在的漏洞信息，主动“引狼入室”，当入侵者上当后按照虚假的指纹信息进行相关攻击，系统可以及时发现，达到早期主动发现入侵的目的。
- 蜜罐防御。网络通信中攻击与防御的问题可视为博弈问题，可在传统蜜罐基础上通过使用模拟服务环境的保护色机制和模拟蜜罐特征的警戒色机制这些主动欺骗技术，使攻击者无法区分蜜罐和实际生产系统，从而达到对攻击者的有效迷惑和诱骗[16-17]。蜜罐的保护色技术是指蜜罐通过模仿周边运行环境和拟保护的生产系统特征，使攻击者无法识别蜜罐的存在。蜜罐的警戒色机制则是指生产系统模仿蜜罐，使得攻击者将系统识别为蜜罐而躲避攻击。

蜜罐防护是攻防双方参与的理性、非合作的诱骗过程，双方策略相互依存，都期望保护自身信息并获得对方信息以达到收益最大化，是一种非合作不完全信息动态博弈[18]。从攻击者视角看，对手不只是提供真实服务的生产系统，而是“蜜罐”和“伪蜜罐”；从防御者视角看，对手则包含合法用户和攻击者。

（3）利用动态的异构冗余机制主动感知网络的攻击

使用多个异构冗余的功能一致体同时运行同一请求，对响应结果进行择多表决，输出正确结果，及时发现网络攻击踪迹，报警异常信息。主动防御网络的动态异构冗余机制可以在不影响系统正常运行的情况下，高准确率下快速发现被入侵部件，并进行系统清洗和恢复。

动态的异构冗余主动防御机制需要解决两个关键问题：

- 异构冗余部件资源池的构建。为了便于进行系统清洗和恢复，可利用虚拟计算技术来构建异构冗余部件资源池[19]。资源池的部件提供相同的服务，但应用程序、操作系统、硬件等需存在差异，以减小异构平台服务器共模故障发生的可能性。
- 部件的选择调度及异常部件的清洗与恢复。可利用综合管控平台按照预定策略完成虚拟机池中虚拟机启动、清洗等调度工作，并且根据管控中心下发的异常部件服务信息执行查杀清洗。周期性或基于事件驱动等策略调度异构部件服务，调度标准是保证系统的功能集不变和尽可能减小系统漏洞交集的关键。

3 主被动防御的组合联动

主动防御机制不是摒弃目前的被动防御系统，不是不需要目前已有的被动防御技术和相关基础设施，而是对目前被动防御系统的拓展、深入和提升，两者之间绝不矛盾，而是相辅相成[20]。

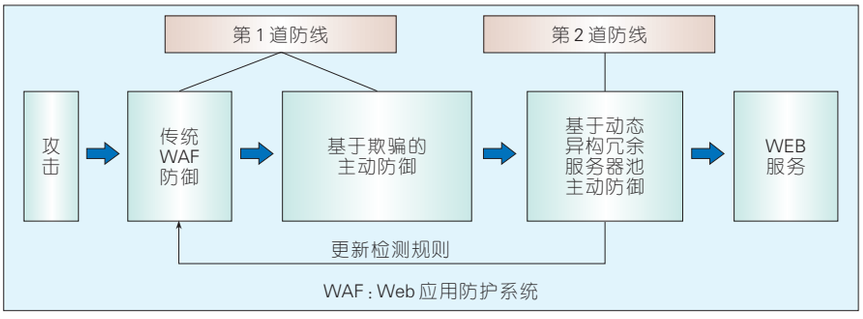
通常来说，可以利用传统的被动防御作为第1道防御阵线，解决大部分目前已知的网络攻击手段的防御问题；利用主动防御作为第2道防

线，解决未知漏洞和后门的防御问题，当然也可部署在第1道防线中。在主动防御发现入侵攻击时，可通过所记录的入侵攻击轨迹进行学习，得到新入侵攻击的特征，对被动防御的特征库和检测规则进行智能更新。被动防御可以利用现有的高效检测机制在入侵到达第2道阵线前过滤掉大部分攻击，主动防御则有效检测第1道防线无法防御的未知攻击，即挡住第1道防线的“漏网之鱼[21]”。

目前，在已开展的面向web应用的主动防御关键技术研究中，我们对主被动联合协作防御技术进行了有效尝试。如图1所示，首先利用已有的Web应用防护系统（WAF）防御大多数已知攻击，利用欺骗伪装技术实现指纹伪装、统一资源定位器（URL）跳变、虚拟蜜罐欺骗、敏感信息过滤、页面信息加扰和头部字段混淆，再利用动态异构冗余机制实现异构冗余体的动态调度、攻击入侵的主动感知和异常部件的有效清洗与恢复。通过上述主被动联合协作防御，不仅实现了web应用已知漏洞防御，而且通过动态变化、欺骗与清洗，可将多个静态的“带毒含菌”web服务应用进行结合，形成了一个动态随机的安全web服务系统。

4 结束语

在当前经济发展的大环境下，如何在保障中国经济快速发展同时做到信息安全的有效防护，是当前学术界与产业界亟需共同解决的一个重大课题。动态网络主动安全防御是解决信息系统中未知漏洞与后门攻击的有效手段，在主动防御系统中，我们不惧怕部件的“带毒含菌”，做到系统总体的安全风险主动可控，可缓解自主产业能力不足的困境，对于改变中国目前甚至今后相当长时期内，特别是在自主可控领域面临的严峻形势下，中国网络安全防御的被动局面具有重要的战略意义和巨大的应用价值。文章介绍了对动态网络主动安全防御技术的若干思考和一些工作，目前中国对相关技术的研究处于起步阶段，所做理论基础研究和相关实践并不充分，动态网络主动安全防御技术的研究任重而道远。



▲图1 主被动联合协作防御的web服务解决方案

参考文献
[1] 邬江兴, 张帆, 罗兴国, 等. 拟态计算与拟态安全防御[J]. 计算机学会通讯, 2015, 11(1): 8–14
[2] 邬江兴. 网络空间拟态安全防御[J]. 保密科学技术, 2014, 10(1): 4–9
[3] 邬江兴. 专题导读——拟态计算与拟态安全防御的原意和愿景[J]. 电信科学, 2014, 30(7): 1–7
[4] 马明杰, 孙奉刚, 翟立东, 等. 网络安全新威胁下我国面临的安全挑战和对策建议[J]. 电信科学, 2014, 30(7): 8–12
[5] COLBAUGH R, GLASS K. Proactive Defense for Evolving Cyber Threats [C]//2011 IEEE International Conference on Intelligence and Security Informatics (ISI), Beijing, China, 2011: 125–130
[6] 姜伟, 方滨兴, 田志宏, 等. 基于攻防随机博弈模型的防御策略选取研究[J]. 计算机研究与发展, 2015, 47(10): 1714–1723
[7] 梁宁宁, 兰巨龙, 程国振, 等. 基于拍卖博弈的可重构服务承载网动态构建算法 [J]. 电信科学, 2015, 31(5): 1–6. doi: 10.11959/j.issn.1000–0801.2015106
[8] CHUANG I, SU W T, KUO Y H. Secure Dynamic Routing Protocols Based on Cross–Layer Network Security Evaluation[C]// 2012 IEEE 23rd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), Sydney, NSW, 2012: 303–308.doi: 10.1109/PIMRC.2012.6362800
[9] 田永春, 刘杰, 隋天宇. 军用无线网络动态安全防御技术研究[J]. 通信技术, 2015, 48(7): 830–834
[10] BALASUBRAMANIAM S, LEIBNITZ K, LIO P, et al. Biological Principles for Future Internet Architecture Design [J]. Communications Magazine, IEEE, 2011, 49 (7): 44–52. doi: 10.1109/MCOM.2011.5936154
[11] ZHOU H, WU C, JIANG M, et al. Evolving Defense Mechanism for Future Network Security [J]. Communications Magazine, IEEE, 2015, 53(4): 45–51. doi: 10.1109/MCOM.2015.7081074
[12] 韦勇, 连一峰, 冯登国. 基于信息融合的网络安全态势评估模型[J]. 计算机研究与发展, 2015, 46(3): 353–362
[13] ALA’ DARABSEH M A A, JARARWEH Y, BENKHELIFA E, et al. SDSecurity: A Software Defined Security Experimental Framework [C]// Third Workshop on Cloud Computing Systems, Networks, and Applications (CCSNA– 2015), London, UK, 2015
[14] SHI L, JIANG L, LIU D, et al. Mimicry Honeypots: A Brief Introduction [C] // Networking and Mobile Computing (WiCOM), 2012 8th International Conference on Wireless Communications, Shanghai, China, 2012: 1– 4. doi: 10.1109/WiCOM.2012.6478572
[15] 付钰, 陈永强, 吴晓平, 等. 基于随机博弈模型的网络攻防策略选取[J]. 北京邮电大学学报, 2014, 37(s1): 35–39
[16] 石乐义, 姜蓝蓝, 刘昕, 等. 拟态式蜜罐诱骗特性的博弈理论分析[J]. 电子与信息学报, 2013, 35(5): 1063–1068
[17] 石乐义, 姜蓝蓝, 贾春福, 等. 拟态式蜜罐诱骗特性的博弈理论分析[J]. 电子与信息学报, 2013, 35(5): 1420–1424
[18] FEINBERG Y. Strategic Communication [C]// Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge, 2011: 1–11
[19] 张翔, 霍志刚, 马捷, 等. 虚拟机快速全系统在线迁移[J]. 计算机研究与发展, 2015, 49(3): 661–668
[20] LIN F Y S, WANG Y S, HUANG M Y. Effective Proactive and Reactive Defense Strategies Against Malicious Attacks in a Virtualized Honeynet [J]. Journal of Applied Mathematics, 2013
[21] 陈锋, 刘德辉, 张怡, 等. 基于威胁传播模型的层次化网络安全评估方法[J]. 计算机研究与发展, 2015, 48(6): 945–954

作者简介

吴春明，浙江大学计算机系统结构与网络安全研究所教授、博士生导师，国家“十二五”信息领域网络与通信技术主题专家组成员；主要研究方向为互联网体系结构、柔性可重构网络、网络资源弹性管控与虚拟化、网络试验床、网络安全主动防御等；曾主持、参加二十余项“973”、“863”、国家自然科学基金、国家科技基础条件平台等项目的研发工作；已发表SCI/EI论文80余篇，授权及申请国家发明专利20余项，出版著作2部。

DOI:10.3969/j.issn.1009-6868.2016.01.010
网络出版地址:http://www.cnki.net/kcms/detail/34.1228.TN.20160112.1457.002.html

编者按:网络空间安全作为一项新的全球治理议程,已经成为世界关注的焦点、各国政府的战略目标之一,但人们对网络空间安全的研究,还缺乏全面系统的理论指导,针对该问题,我刊特转载自《科学网》一篇由北京邮电大学杨义先、钮心忻教授编写的《安全通论》(原文网址:http://blog.sciencenet.cn/blog-453322-944217.html)。在该文章中,作者提出需要建立一套基础的通用安全理论,来指导包括网络空间安全在内的所有安全保障工作;并从安全角度出发,形象地将系统比作完整的"经络树",认为对系统的任何"病痛"都可进行有效的"医治"。

安全通论——经络篇

The General Theory of Security: Meridian

中图分类号:TN929.5 文献标志码:A 文章编号:1009-6868 (2016) 01-0038-004

摘要:从安全角度出发,用概率方法严格证明了任何有限系统,都存在一套完整的"经络树",使得系统的任何"病痛",都可以按如下思路进行有效"医治":首先梳理出经络树中"受感染"的"带病树枝"体系,然后对该树枝末梢上的"带病树叶"("穴位"或"元诱因")进行"针灸",医治好病叶后,与这些病叶相连的树枝就治好了;医治好所有病枝后,与这些病枝相连的"树干"就治好了;医治好所有"病干"后,整棵经络树就医治好了,从而系统的病痛就治好了。此处所指的有限系统,既可以是儿童玩具这样的微系统,也可以是芯片、计算机、电信网、互联网、物联网甚至整个赛博空间等复杂巨型有限系统;当然也可以是消防、抗灾、防病、治安、环保等各类常见的其他系统。

关键词:网络空间安全;国家安全;健康防御;人体经络

Abstract: From a security point of view, and proved by probability method, we point out that in any finite system there are a complete set of " meridian trees" that makes any "pain" effectively cured according to the following ideas: Firstly, we need to find out the "infected branch" system, then do the "acupuncture" to the "sick leaves (acupoint) " at the end of the bunch. After curing the sick leaves, the branch that is connected with these leaves is cured; after curing the sick leaves, the "trunk" that is connected with these branches is cured; after curing all the "sick trunks", the whole tree is cured, so the pain of system is cured. The system here refers to the finite system, and it can be both a micro system, and also a chip, computer, telecommunication network, Internet, Internet of things or even the cyber space of complex giant system; of course, it can also be a fire, disaster, disease prevention, public security, environmental protection and other types of common other systems.

Key words: network space security; national security; health defense; human meridian

杨义先/YANG Yixian
钮心忻/NIU Xinxin
(北京邮电大学 信息安全中心,北京 100876)
(Information Security Center, Beijing University of Post and Telecommunication, Beijing 100876, China)

- 安全是一个与角度、时间、对象都密切相关的概念
- 不安全性遵从热力学第二定律
- 经络图中平均概率值大的"经络"是更脆弱的经络,是在系统安全保障中需要重点保护的部分,也是攻击过程中重点打击的部分
- 绘制出网络空间的安全经络图需要投入很多的精力并耗费很长的时间

安全与信息都是至今还没有严格定义的概念,但是这并不意味着不能对它们进行深入研究,其实早在60年前,仙农就已经创立了信息论,从而为现代通信的飞速发展奠定了坚实的基础。但是,至今人们对安全的研究,特别是网络空间安全的研究,还仅仅停留在"兵来将挡,水来土淹"的工程层次或技术层次,既缺乏全面系统的理论指导,又遗留了许多明显的漏洞,比如虽然大家都承认网络空间安全是"三分技术,七分管理",但全世界都将几乎90%的精力

收稿日期:2016-01-04
网络出版时间:2016-01-12

聚焦于那三分技术，七分管理竟然无人问津，或者说人们只是片面地将管理理解为“颁布几份规章制度”而已。

我们梦想建立一套基础的通用安全理论，并以此来指导包括网络空间安全在内的所有安全保障工作。文章是努力实现该梦想的第1步，希望能够激发学者们更多的后续研究。

1 不安全事件的素分解

安全是一个很主观的概念，与角度密切相关。同一个事件，对不同的人，从不同的角度来说可能会得出完全相反的安全结论，比如，政府监听公民通信这件事，从政府角度来看，能监听就是安全；而对公民来说，能监听就是不安全。所以，我们研究安全，只锁定一个角度，比如，“我”的角度。

安全是一个与时间密切相关的概念。同一个系统，在昨天安全，绝不等于今天也安全（比如，若用现代计算机去破译古代密码，简直是易如反掌）；同样，在今天安全，也绝不等于明天就安全。当然，一个在昨天不安全的系统，今天也不会自动变为安全。因此，我们研究安全时，只考虑时间正序流动的情况，即立足当前，展望未来。

安全是一个与对象密切相关的概念。若A和B是两个相互独立的系统，若我们只考虑A系统的安全，B系统的安全就应该完全忽略。因此，我们研究安全时，只锁定一个有限系统，即该系统由有限个“元件”组成。

假设A是一个封闭的独立系统，如果直接研究其安全，根本就无处下手！不过，幸好有“安全”=不“不安全”，所以，若能把不安全研究清楚了，安全也就明白了。

假设A系统中发生了某个事件，如果它是一个对“我”来说的不安全事件，那么“我”就能够精确且权威地判断这是一个不安全的事件，因为该事件的后果是“我”不愿意接受的！需要注意：除“我”之外，“别人”的判断是没有参考价值的，因为，文中只从一个角度来研究安全。如果将该不安全事件记为D，该事件导致系统A不安全的概率就记为$P(D)$。我们只考虑$0<P(D)<1$的情况，因为如果$P(D)=0$，这个不安全事件就几乎不会发生，故可以忽略，因为无论是否对造成事件D的环境进行改进，都不影响系统A的安全性；如果$P(D)=1$，D则是不安全的确定原因，这时只需要针对事件D单独进行加固，就可以提升系统A的安全性了。

从理论上看，给定系统A之后，如果A是有限系统，则可以通过各种手段，发现或测试出当前的全部有限个不安全事件，比如，D_1、D_2、……、D_n。在不引起混淆的情况下，我们用D_i同时表示不安全事件和造成该事件D_i的原因。于是，系统A的不安全概率就等于$P(D_1\cup D_2\cup...\cup D_n)$，或者说，系统A的安全概率等于$1-P(D_1\cup D_2\cup...\cup D_n)$。

换句话说，本来无处下手的安全研究，就转化为了安全数学问题，即在概率$0<P(D_1\cup D_2\cup...\cup D_n)<1$的情况下，使该概率$P(D_1\cup D_2\cup...\cup D_n)$最小化的问题，或者使$1-P(D_1\cup D_2\cup...\cup D_n)$最大化的问题。

假设D和B是系统A的两个不安全事件，$(D\cup B)$则也是一个不安全事件，但是$(D\cap B)$或者$(D\backslash B)$等就不一定再是不安全事件了。如果事件D是B的真子集，并且D的发生会促使B也发生，则称事件D是事件B的“子事件”。

在时间正序流动的条件下，假设系统A的过去全部不安全事件集合为D，若当前又发现一个新的不安全事件B，则有系统A的当前不安全概率$=P(D\cup B)\geqslant P(D)$=系统A的过去不安全概率。于是，不安全性遵从热力学第二定律：系统A的不安全概率将越来越大，而不会越来越小（除非有外力，比如采取了相应的安全加固措施等）；或者说安全与信息一样都是负熵。

假设Z是一个不安全事件，如果存在另外两个不安全事件X和Y（它们都是Z的真子集），同时满足如下两个条件：$X\cap Y=\varnothing$（空集）；$Z=X\cup Y$，我们就认为不安全事件Z是可分解的。此时X和Y都是Z的子事件。如果某个不安全事件是不可分解的，即它的所有真子集都不再是不安全事件了，我就称该事件为不安全的素事件。

定理1（不安全事件分解定理）：对任意给定的不安全事件D，都可以判断出D是否可分解，如果是可分解的，也可以找到它的某种分解。

证明：由于有限系统A的全部不安全事件只有有限个，即D_1、D_2、……、D_n，所以至少可以通过穷举法，对每个$D_i(i=1,2,\cdots,n)$测试一下$D\backslash D_i$，看看它是否也是不安全事件。如果至少能够找到某个这样的i，那么D就是可分解的，而且D_i与$(D\backslash D_i)$就是它的一个分解；否则，如果这样的i不存在，那么D就是不可分解的不安全素事件，这是因为$D_1,D_2,\cdots,D_n$是全部不安全事件。证毕。

定理2（不安全事件素分解定理）：若反复使用上述的不安全事件分解定理来处理不安全事件$(D_1\cup D_2\cup...\cup D_n)$及其被分解后的不安全子事件，那么就可最终得到分解$D_1\cup D_2\cup...\cup D_n=B_1\cup B_2\cup...\cup B_m$，这里对任意的$i$和$j$（$i,j=1,2,\cdots,m$）都有$B_i$是不安全素事件并且$B_i\cap B_j=\varnothing$（空集）。

证明：若$D=D_1\cup D_2\cup...\cup D_n$已经是不可分解的了，则有$m=1$，并且$D_1\cup D_2\cup...\cup D_n=B_1$。

如果D是可以分解的，并且X是D分解后的一个不安全子事件。如果X已经不可分解了，则可以取$B_1=X$；如果X还可以再分解，再对X的某个不安全子事件进行分解。如此反复，直到最终找到一个不能再被分解的不安全子事件B_1。

仿照上面分解D的过程，来试图分解$D\backslash B_1$，便可以找出不能再被分解

的不安全子事件B_2。再根据$D\backslash B_1\cup B_2$的分解，便可得到B_3。

最终，当这个分解过程结束后，全部的B_i就已经构造出来了。证毕。

于是，根据不安全事件素分解定理，便有$B_i\cap B_j=\emptyset$（空集），并得出$P(D_1\cup D_2\cup...\cup D_n)=P(B_1\cup B_2\cup...\cup B_m)=P(B_1)+P(B_2)+...P(B_m)$，因此换句话说，我们可以将引发有限系统A的不安全事件D_1、D_2、……、D_n，分解为另一批彼此互不相容的不安全素事件B_1、B_2、…、B_m，并且，还将有限系统A的不安全概率转化为$P(B_1)+P(B_2)+...+P(B_m)$。所以，有限系统A的不安全概率$P(D_1\cup D_2\cup...\cup D_n)$的最小化问题，也就转化成了每个彼此互不相容的不安全素事件的概率$P(B_i)$ $(i=1,2,\cdots,m)$的最小化问题。

定理3（分而治之定理）：任何有限系统A的不安全事件集合，都可以分解成若干个彼此互不相容的不安全素事件：B_1、B_2、…、B_m。使得只需要对每个$B_i(i=1,2,\cdots,m)$进行独立加固，即减小事件B_i发生的概率$P(B_i)$，就可以整体上提高系统A的安全强度，或者说整体上减少系统A的不安全概率。

分而治之定理回答了前面的热平衡问题，即有限系统A的不安全状态，将最终稳定成一些彼此互不相容的不安全素事件之并。该定理对全球网络空间安全界的启发意义在于：过去那种“头痛医头，足痛医足”的做法虽然值得改进，但也不能盲目地“头痛医足”或“足痛医头”，而是应该科学地将所有安全威胁因素，分解成互不相容的一些“专科”（B_1、B_2、…、B_m），然后，再开设若干“专科医院”来集中精力“医治”相应的病症（即减小$P(B_i)$）。

2 系统“经络图”的逻辑分解

设X是B的一个真子集，如果事件X发生，将促进B也发生（即$P(B|X)-P(B)>0$），我们就称X为B的一个诱因。

针对任何具体给定的有限系统A，因为B是有限集，所以从理论上看，总可以通过各种手段发现或测试出当前B的全部有限个诱因，比如，X_1、X_2、…、X_n，即$B=X_1\cup X_2\cup...\cup X_n$。

设X和Y是B的两个诱因，而且还同时满足：$X\cap Y=\emptyset$（空集）；$B=X\cup Y$。我们则认为B是可分解的，并且$X\cup Y$就是它的一种分解。如果某个B是不可分解的（即它的所有真子集都不再是其诱因了，或者说对B的所有真子集Z，都有条件概率$P(B|Z)=P(B)$），我们就称该事件为素事件。

若Y、Y_1、Y_2都是B的诱因，并且$Y_1\cap Y_2=\emptyset$（空集）；$Y=Y_1\cup Y_2$，我们则认为B的诱因Y是可分解的，并且$Y_1\cup Y_2$就是它的一种分解。如果诱因Y是不可分解的（即它的所有真子集都不再是B的诱因了），我们就称该诱因Y为B的素诱因。如果诱因Y的所有子集Z，都不再是Y自己的诱因了，我们就称Y为元诱因，或形象地称为“穴位”。

定理4（事件分解定理）：对任意给定的事件B，都可以判断出其是否是可分解的，如果是可分解的，也可以找到它的某种分解。

证明：由于系统B的全部诱因只有有限个，即X_1、X_2、……、X_n，所以至少可以通过穷举法，对每个$X_i(i=1,2,\cdots,n)$测试一下$B\backslash X_i$，看看它是否也是B的一个诱因。如果至少能够找到某个这样的i，那么B就是可分解的，而且X_i与$(B\backslash X_i)$就是它的一个分解；如果这样的i不存在，那么B就是不可分解的，这是因为X_1、X_2、……、X_n是B的全部诱因。证毕。

定理5（事件素分解定理）：若反复使用上述的事件分解定理来处理事件B，就可以最终得到分解，即$B=Y_1\cup Y_2\cup...Y_m$，这里对任意的i和j（$i,j=1,2,\cdots,m$）都有$Y_i\cap Y_j=\emptyset$（空集），并且每个Y_i都是B的素诱因。

证明：若B已经是不可分解的了，则有$m=1$，$B=Y_1$。

假设B是可以分解的，且Y是B分解后的一个诱因。如果Y已经是B的素诱因了，则可以取$Y_1=Y$；如果Y还可以再分解，则再对Y的某个诱因进行分解。如此反复，直到最终找到一个不能再被分解的素诱因，请将它记为Y_1。

仿照上面分解B的过程，来试图分解$B\backslash Y_1$，便可以找出B的不能再被分解的素诱因Y_2。

再根据$B\backslash(Y_1\cup Y_2)$的分解，便可得到Y_3。

最终，当这个分解过程结束后，全部的Y_i就已经构造出来了。证毕。

有了上面各定理的准备后，我们现在就可以给出如下的有限系统A的经络图算法步骤。

第0步：针对系统A的不安全事件D。

第1步：利用定理2，将D分解成一些互不相容的不安全素事件$B_1\cup B_2\cup...\cup B_m$，这里对任意的$i$和$j$（$i,j=1,2,\cdots,m$）都有$B_i$是不安全素事件并且$B_i\cap B_j=\emptyset$（空集）。在绘制经络图时，可以从左至右，按照$P(B_i)$的递减顺序排列。

第2.i步（$i=1,2,\cdots,m$）：利用定理5，把第1步中所得到的B_i分解成若干B_i的素诱因，在绘制经络图时，可以从左至右，对B_i的素诱因，按照其发生概率大小值的递减顺序排列。为避免混淆，我们将所有第2步获得的素诱因，称为第2步素诱因。这些素诱因中，有些可能已经是元诱因（穴位）了。

第3.i步（$i=1,2,\cdots,m$）：针对第2步所获得的每个不是元诱因（穴位）的素诱因，利用定理5，将其进行分解，由此得到的素诱因，称为第3步素诱因（这些诱因的从左到右的排列顺序也与前几步相似）。这些素诱因中，有些可能已经是元诱因（穴位）了。

……

第$k.i$步（$i=1,2,\cdots,m$）：针对第$k-1$步所获得的不是元诱因（穴位）的

每个素诱因，利用定理5，将其进行分解，由此得到的素诱因，称为第k步素诱因（这些诱因的从左到右的排列顺序也与前几步相似）。这些素诱因中，有些可能已经是元诱因了。

由于上面各步骤的每次分解，都是针对真子集进行的，所以这种分解的步骤不会无穷进行下去，即一定存在某个正整数，比如N，使得在第$N.i$步（i=1，2，…，m）中，针对第N−1步所获得的不是元诱因的每个素诱因，利用定理5，将其进行分解，由此得到的素诱因全部都已经是元诱因（穴位）了（每一个素诱因下面的元诱因排列顺序，也是采用了概率从大到小进行）。

将上面的分解步骤结果，用图形表述出来，我们便得到了有限系统A的不安事件“经络图”，由于它的外形很像一棵倒立的树，所以我们也称这为“经络树”，如图1所示。

现在我们就比较清楚，该如何头痛医足了：实际上，只要系统A“病”了，就一定能够从系统A的完整经络图中找出某个“生病的子经络图”M，使得（1）M的每层素诱因或元诱因（穴位）都是病的；（2）除了M之外，系统A的经络图的其他部分都没病。为了治好该病，只需要将M中的所有元诱因（穴位）的病治好就行了，即只需要对这些元诱因（穴位）扎针灸就行了。（说明：这里某个第k步诱因病了，指它的至少一个第k+1步诱因发生了；而如果某个第k步诱因的全部第k+1步诱因都没有发生，那么这个第k步诱因就没病！除了元诱因（穴位）之外，M中的其它非元诱因是可以自愈的！）

更具体地说，头痛医足的过程是：首先将最底层，比如第N层的元诱因（穴位）治好，第N−1层的素诱因就自愈了；然后，再扎针灸治好第N−1层的元诱因（穴位），第N−2层的素诱因就自愈了；然后，再扎针灸治好第N−3层的元诱因（穴位），如此继续，最终到达顶层，就可以了。

经络图的用途显然不仅仅是用来头痛医足，它还有许多其他重要应用，比如：

（1）只要守住所有相关的元诱因（穴位），系统A就安然无恙。

（2）同理，只要所有炮火瞄准相关元诱因（穴位），那么就能够稳准狠地打击对手。

（3）除了元诱因（穴位）之外，经络图中平均概率值大的“经络”是更脆弱的经络（即安全“木桶原理”中的短板），也是在系统安全保障中需要重点保护的部分，同时也是攻击过程中重点打击的部分。

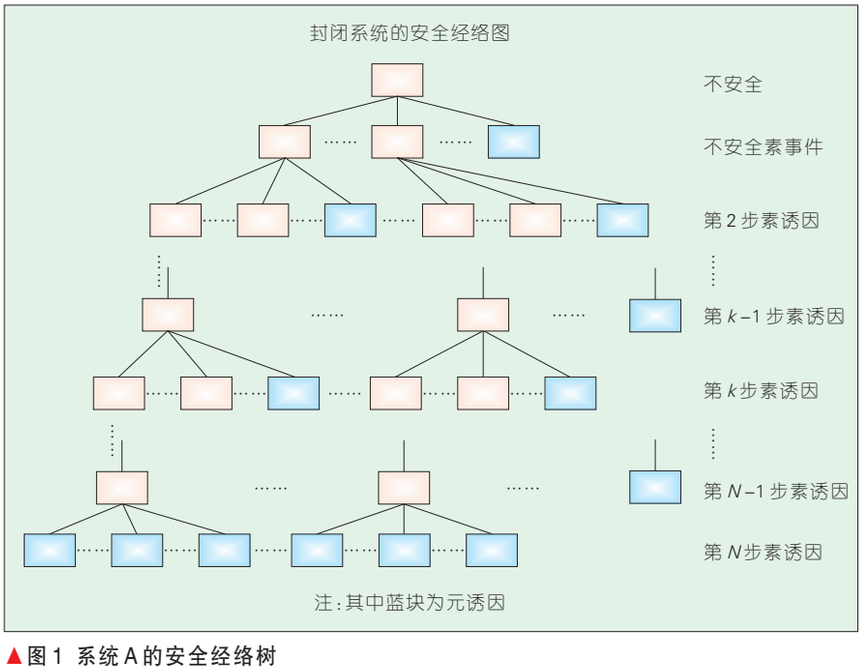
（4）平时就可绘制和补充经络图，在关键时刻就可以排上用场了！

3 结束语

仙农在研究信息论时，虽然发现了信道容量的上限值，但是他没能给出如何才能达到该上限值，从而使全世界通信界的科学家们在过去60余年里，设计各种编码方法来努力逼近仙农界，至今没有成功。

与此相似，文章中虽然证明了有限系统的安全经络图是存在的，但是并未给出如何针对具体的系统，来绘制其安全经络图。估计未来的学者们也不得不花费巨大的精力，针对具体系统来绘制具体的经络图。

必须指出：绘制经络图绝非易事。想想看，为了绘制人体经络图，中医界的祖先们奋斗了数千年！如今我们也需要很长时间才能绘制出网络空间安全经络图。



▲图1 系统A的安全经络树

作者简介

杨义先，灾备技术国家工程实验室主任，北京邮电大学教授、博士生导师，信息安全中心主任，首批长江学者特聘教授，首届国家杰出青年基金获得者，中国密码学会副理事长；目前研究方向为网络空间安全、现代密码学和纠错编码等；获得包括国家发明奖和省部级科技进步奖等在内的各类科技奖励20余项，授权发明专利4项，主持和参与多项国家“863”、国家自然科学基金、省部级等科研项目；发表高水平论文500余篇，出版专著及教材20多部。

钮心忻，北京邮电大学计算机学院教授、博士生导师；长期从事网络与信息安全、信号与信息处理等方面的研究工作。

DOI:10.3969/j.issn.1009-6868.2016.01.011
网络出版地址:http://www.cnki.net/kcms/detail/34.1228.TN.20151224.1650.006.html

信息设备供电系统发展趋势

Development Trend of Power System for ICT Equipment

胡先红/HU Xianhong
(中兴通讯股份有限公司能源产品部,广东 深圳 518057)
(Energy Product Department, ZTE Corporation, Shenzhen 518057 , China)

中图分类号:TN929.5 文献标志码:A 文章编号:1009-6868 (2016) 01-0042-004

摘要: 认为降低总体拥有成本(TCO)是信息设备供电系统发展的最直接驱动力,高效、节能、绿色、共享、智能、互联是未来信息设备供电系统的发展趋势和特征。能源互联网可以实现能源的互联共享,减少冗余浪费,提高供电可靠性,是未来各种能源系统的发展方向。信息设备供电系统是智能化最高的能源设备,可能是最先成为能源互联网"终端"的设备。

关键词: 信息设备供电;高效;节能;绿色;共享;智能;互联;能源互联网

Abstract: In this paper, we propose that ICT power system is driven by the lower total cost of ownership (TCO). The development trends of ICT power systems are: high efficiency, energy conservation, green, sharing, intelligent, and interconnection. Energy Internet can realize the interconnection of energy, reduce redundant waste, and increase power reliability. It is the future development direction of various energy systems. ICT power system is the highest intelligent energy equipment and will probably first become a terminal of energy internet.

Keywords: ICT power supply; high efficiency; energy conservation; green; sharing; intelligent; interconnection; energy Internet

在M-ICT时代,连接无处不在。高度连接的信息社会,对能源的使用要求越来越高,ICT系统消耗的能源也不断增加,绿色节能成为M-ICT时代的迫切需求,也越来越得到人们的重视。另一方面,新能源技术的发展也为绿色节能的信息设备供电提供了有力的支撑。信息设备供电系统呈现出高效、节能、绿色、共享、智能、互联等特征。

1 高效

高效是指在ICT设备供电系统中,从能源的产生、转换、分配、使用等能效矩阵的各个环节,对能源进行高效的转换和利用,各个环节的高效汇集成整个供电系统的高效。高效意味着在供电系统有着更小的能量损失。

1.1 能源的高效率转换

ICT设备所直接使用的是5 V、3.3 V等低压直流,从能源的供给到最终的低压直流电源之间,有很多的能源转换、分配环节。每一个转换分配环节,都存在能量的损失。高效率的电源变换技术,使得能源在各个转换环节的损失最小。各种功率变换模块都在追求更高的转换效率。

(1)通信用电源整流器的(交流220 V转直流48 V)效率,已从92%提升到当前主流的96%以上,未来将进一步提升到98%甚至更高。从92%到96%,再到98%,整流器损耗依次降低50%。采用高效率的整流器,能有效降低站点的电能消耗。

(2)通信站点用太阳能功率模块(直流到直流(DC/DC))效率提升到98%以上。

(3)数据机房用不间断电源(UPS)效率提升到97%以上。

(4)太阳能极板光电转换效率不断提升,例如单晶硅的量产效率达到20%以上,并且每年提升绝对值0.3%~0.5%左右。光电转换效率的提升可以提高太阳光的利用率,产生更多的电能。

(5)各种能源变换器中轻载下的效率不断提升,满足实际负载情况下的节能,通信整流器的最佳效率点出现在50%~80%负载,30%以上负载就能实现很高的效率。

1.2 高效率的能源利用

信息设备供电除了来源于公共电网,还来自于新能源发电、油机发电等。为了降低成本,这些小型发电设备的配置容量都非常有限,因此能源的高效利用则成为了很重要的研究内容。

(1)最大功率点跟踪(MPPT)技术,成为太阳能、风能利用中的普遍需求。太阳能功率模块MPPT效率达到99%,通过实时跟踪太阳能或风机

收稿日期:2015-11-20
网络出版时间:2015-12-23

的最大输出功率点，可以实现最大功率的利用，可比传统控制器增加30%的利用率。

（2）油机发电的变频输出，可以提高轻载燃油效率。传统交流油机，工作在恒定频率状态，在轻载状态下，燃油效率较低。通过采用新型直流油机，油机发电机工作在变频状态，输出频率随负载调整，使得燃油效率有较大提升。

1.3 高效率的供电架构

不同的供电架构有着不同的能源效率。数据流量的指数级增长，推动了大量的数据中心的建设，也催生了新型供电架构的应用，供电架构的探讨成为互联网数据中心（IDC）机房供电领域的热门话题。

（1）高压直流取代交流UPS供电成为趋势[1-2]。高压直流（HVDC）供电比传统交流UPS减少了输出直流电/交流电（DC/AC）逆变、负载侧的整流环节等，供电转换环节少，供电效率高。电池直接挂在母线上，可靠性更高。行业内通用的高压直流有240 V和336 V两种制式之争，240 V制式对目前的交流供电系统兼容性好，主要在中国国内应用；336 V则更为彻底，能够节省更多转换的环节，在全球范围内得到广泛认可。无论哪种制式最终胜出，都会逐步取代交流UPS而成为数据机房供电的主流，并带来供电效率的提升。

（2）市电直供技术逐步推广。为进一步提升供电效率，减少转换环节，市电直供主用、高压直流备用的供电架构得到越来越多的应用。这种架构在正常情况下采用市电直接供电，进一步提升了供电的效率[3]。

（3）中间总线电压存在争论。传统服务器采用12 V的中间母线供电架构，而传统通信设备则采用48 V为供电母线电压。随着服务器的功耗越来越大，母线电阻的损耗不可忽视，于是出现了采用48 V作为中间母线的供电架构，可以大幅降低母线电流。而48 V基本上是人可接触的安全电压上限，较好地实现了效率与安全之间的平衡[4-5]。

2 节能

节能主要是指通过各种节能措施，在保证信息业务可靠性的基础上，降低信息设备站点和机房的能源消耗。随着节能减排的要求不断提高，对已部署的设备实行节能改造，往往能够有效降低站点能耗，并降低站点的运营成本（OPEX）。

2.1 休眠技术节能

在信息设备供电系统中，大多数都采用模块化并联实现冗余以保证可靠性，并且为保证最大负载情况的供电，系统设备配置都偏大。正常情况下的负载率一般在20% ~ 40%，并不在电源设备的最佳工作状态。模块休眠可以根据负载的情况关闭或开启部分模块，使剩余模块以50% ~ 70%负载率工作在最佳效率状态，并让系统的各模块轮换进行休眠已使得各模块工作时间均衡。通信直流电源系统中的整流器休眠技术可以实现节电4% ~ 10%。模块化UPS、高压直流供电系统均可采用模块休眠技术实现节能。在负载侧，也实现了根据用户量实现设备的休眠，大幅减少信息设备自身的功耗。

2.2 混合供电节能

在偏远无市电通信站点，为实现通信，一般采用双油机轮换供电，蓄电池浮充备用。在这种情况下，油机配置都有较大的冗余量以适应电池充电、未来扩容等，从而导致正常情况下油机带载率低、燃油效率低。通过改造，可以实现油机、电池轮流混合供电，例如油机工作5 h，给负载供电的同时给电池充电；然后油机停机，电池放电给负载供电7 h；然后油机再工作，如此循环。采用这种方式，油机负载率大幅提升，燃油效率有效提高，油机工作时间减少，燃油消耗减少，油机维护费用也下降[6]。

2.3 热管理节能

在传统的信息设备机房，制冷消耗了大约40%电能。一方面，我们可以采用高效、节能的信息设备供电系统，以减少机房设备的热量产生；另一方面，可以通过新型的热管理技术，有效降低机房的制冷耗能[7]。这些新型的热管理技术包括：

（1）机房精确制冷。即改变机房整体制冷的情况，将冷风精确送达需要冷却的部件，实现精确制冷，并通过空间布局设计防止冷热风短路，这样可以有效降低制冷空调的耗电量。

（2）分区制冷。在信息设备机房中，不同设备对温度的要求不同。通过对不同的区域实现分区制冷策略，可以减少制冷的能耗。例如，对电池单独实现温控，同时适当调高机房的整体温度，可以有效减少空调耗能。

（3）户外机柜取代机房。对于小型站点，采用户外设备（含电源系统、基站设备）取代户内设备，相比机房，可有效减少制冷空间，大幅度降低制冷能耗，也能减少机房占地面积和建设费用。

（4）采用新风技术、自然冷源等散热技术。采用新风技术，在室外环境温度低、昼夜温差大、空气洁净的地区，可以有效减少空调工作时间。在大型水库、湖泊、海边，可充分利用自然冷源进行机房冷却，降低机房制冷能耗。

（5）电能利用率（PUE）被用来衡量机房的绿色程度。国际先进的数据中心PUE值能达到1.1。在通信行业标准中，对各种机房的PUE值有比较明确的要求[7]。量化的标准将推动机房的能耗降低，PUE值的提高。

3 绿色

绿色是指采用可再生能源、清洁能源取代化石能源为信息设备供电，以减少站点碳排放。在信息设备供电的清洁能源中，主要有太阳能、风

能、燃料电池等。

3.1 太阳能绿色供电

随着太阳能发电的成本近年来大幅度的降低，太阳能供电的经济性大幅提高。在没有市电或市电不稳定的站点，太阳能供电得到了大量的应用；而在市电较好的地方，通过配置一定的太阳能，也能有效减少信息设备对市电的消耗[8]。太阳能供电将是未来信息设备绿色供电的主流。

3.2 风光互补混合供电

风能和太阳能在昼夜、季节上存在一定程度的互补性，例如白天阳光强、风弱，晚上无太阳能、风大，冬季阳光弱、风强。在风能比较好的地区，通过风光互补方式给信息设备站点供电，也得到了较多的应用。

3.3 燃料电池绿色供电

燃料电池在通信站点的清洁供电试点越来越多，但目前燃料电池还是以氢为主要燃料（其他非氢燃料电池也需要最终制备成氢气进入燃料电池），由于成本高、燃料获取的不便利等特点，燃料电池的供电普及受到了比较大的限制。随着技术的成熟和成本的降低，燃料电池将越来越多地应用于信息设备供电。

4 共享

共享能有效节省空间、硬件资源，在信息设备供电系统中，共享在各个层次展开。

4.1 共享部件

不同功率的整流器（如2 000 W与3 000 W）、不同类型的变换器（如太阳能模块与整流器模块）在物理和电气接口实现兼容，可以共享插箱槽位，减少系统槽位空间。同一变换器，可以兼容交流输入、直流输入，以及不同的电压输出。同一电源插箱，可以实现市电、油机、太阳能、风能、电池的接入，统一输出并监控，实现了部件级的共享，减少了重复冗余的硬件资源和空间。

4.2 共享机柜

随着功率模块、监控单元、电池等功率密度不断提高，各种不同单元在单机柜内实现一体化，同时在部分室外站点，信息设备与电源设备共用机柜，实现了整个站点的一体化，减少了设备占地空间。

4.3 共享站点

大多数情况下，通信运营商在同一站点实现了无线2G、3G、4G通信的共享供电。中国铁塔公司的成立，意味着中国国内运营商将全面进入站点共享，包括电源在内的基础设施共享。随着非洲、拉丁美洲的基建运营商的规模不断扩大，全球范围内的通信站点共享越来越多。站点共享能大幅度降低通信运营商在基础设施建设和运营上的费用支出。

随着电动汽车的发展，通信站点与电动汽车充电融合的方案也开始实施，以实现信息设备供电与电动汽车充电的站点共享。

4.4 共享能源

目前物理地址不同的站点之间还不能实现能源的共享，每个站点都需要备用电池、油机等备电设备，每个站点的新能源发电也是站点内部使用。随着能源互联网的发展，未来不同站点共享备电和供电将成为可能。汽车到电网(V2G)技术也会让电动汽车与通信站点共享电池能量。能源的共享可以有效降低各个站点的备用容量，降低设备投资，供电可靠性也得以提高。

5 智能

智能化伴随着信息设备供电系统的发展不断发展，从早期的站点参数的监控扩展到站点的能源数据管理，智能化已经可以有效降低站点运维成本和能源消耗。

（1）智能监控。信息设备供电系统的智能监控功能实现对供电系统的信息采集、监测、告警和远程控制。实现了站点的无人值守，维护成本得以降低。

（2）数字控制技术。数字信号处理（DSP）数字控制技术具有控制灵活、成本低、保密性好、时间稳定性好等特点，在各种功率变换逐渐成为主流。而在复杂的功率变换系统，必须靠DSP数字控制技术才能实现复杂的变换。

（3）能源数据管理。站点的能源数据管理除了能够实现站点的设备监控外，还可以进一步实现站点预防性维护，提醒设备维护人员进行预防维护保养；通过站点能耗分析，可以实现站点能效管理；通过远程读取设备资产信息，可以实现在网资产的自动统计和盘点等资产管理；通过站点门禁、防盗视频抓拍等，可以实现站点安全管理。

（4）软件定义。DSP为软件定义电源奠定了基础。软件定义的电源系统逐步出现，可以通过软件定义电源的特性，实现能量的双向流动。随着功率器件的发展，四象限器件的出现，更可以同一硬件电路，实现AC/DC、DC/AC、DC/DC之间的任意功率变换[9]。

（5）自适应。产品的智能化可以根据使用的外部环境如电网、季节、天气、温度、负载，调整电源系统的工作状态，如输出电压、频率、负载率、电池充电状态等，以使得系统工作在最优状态，减少电能消耗、延长设备寿命。

6 互联

当前的信息设备供电系统，实现了远程的网络监控，已经具备了信息互联。随着新能源的应用普及，每个站点不仅消耗能源，也能产生能源、储存能源。

新能源站点多余的能源可以实现供电的互联，共享给其他不同物理

地址的站点，实现通信站点供电的互联共享，并可进一步进入公用供电网络，成为大的能源互联网的子网。

信息设备供电站点的电池除了可以用来做停电时给站点设备供电外，还可以共享给其他站点，也可以作为能源互联网削峰填谷、平滑输出的储能单元。

能源互联网是信息技术与新能源相结合的产物，虽然还没有一个统一的定义，但智能、开放、互联、共享等是能源互联网的一些主要特征。能源互联网可以实现能源的互联共享、峰谷互补，减少冗余浪费，提高供电可靠性[10]。信息设备供电系统是智能化最高的能源设备，可能是最先成为能源互联网“终端”的设备。

7 结束语

信息设备供电系统的发展往往决定于以下两个因素：一是产品的使用者（用户）的需求变化；二是技术的发展。前者是外因，后者是内因。降低总拥有成本（TCO）是信息设备供电系统发展的最直接的驱动力，高效、节能、绿色、共享、智能、互联最终都能带来信息设备供电系统的TCO降低，各种新材料、新器件、新应用技术为系统的发展提供了技术支撑。能源互联网是未来各种能源系统的发展方向，虽不能准确定义，不能准确预测，但都是朝着这个方向发展着。

节能减排既是企业利益诉求，也是企业的社会责任。我们将致力于推动信息设备供电系统更为高效、节能、绿色、共享、智能、互联，推动信息技术系统的创新发展。

参考文献

[1] MCEACHERN Alex. Energy for Telecommunications: the Next 25 Years [C]// The 37th International Telecommunications Energy Conference (INTELEC2015), Osaka, Japan, 2015
[2] TAKASHI O, SHINTARO O. Strategy for Introduction of High-Voltage DC Power Supply System at NTT Group[C]// The 37th International Telecommunications Energy Conference (INTELEC2015), Osaka, Japan, 2015
[3]彭厂香. 基于336V直流的市电直供技术[J]. UPS应用, 2015(11): 38-43
[4] TORU T, HIROAKI M. The HVDC Power Supply System Implementation in NTT Group and Next Generation Power Supply System [C]// The 36th International Telecommunications Energy Conference (INTELEC2014), Vancouver, Canada, 2014
[5] 李典林. 数据中心未来供电技术发展浅析[J] 电信网技术, 2014(10) : 47-52
[6] LIU M M. A Resilient Hybrid Energy Power System Architecture[C]// The 36th International Telecommunications Energy Conference (INTELEC2014), Vancouver, Canada, 2014
[7] YD/T 3032-2015 通信局站动力和环境能效要求和评测方法[S]
[8] WANG Y, XU L. A New Integrated Hybrid Power Supply System for Telecom Site Sharing Solution[C]// The 37th International Telecommunications Energy Conference (INTELEC2015), Osaka, Japan, 2015
[9] Bruce Carsten. The Past and Future of Power Electronics[C]// The 36th International Telecommunications Energy Conference (INTELEC2014), Vancouver, Canada, 2014
[10] 冯庆东. 能源互联网与智慧能源[M]. 北京: 机械工业出版社, 2015

作者简介

胡先红，中兴通讯股份有限公司能源产品总工程师、能源产品规划首席专家，广东省电源行业协会副会长、中国电源学会标准化工作委员会副主任委员；主要负责中兴通讯能源产品的规划与架构方案设计工作。

综合信息

中兴通讯技术杂志社北京迎春联谊会隆重召开

【本刊讯】2016年1月8日，“中兴通讯技术杂志社2016北京迎春联谊会”在北京新世纪日航饭店隆重召开。刊物在京编委以及来自高等院校、科研院所、通信运营商、期刊管理部门等80多位代表共聚一堂，聊通信热点，话未来发展，气氛轻松而热烈。

中兴通讯股份有限公司高级副总裁、杂志社新任总编陈杰到会并发言。她感谢与会代表对公司及刊物的关心和支持，并介绍了2015年公司在技术、产品、市场等方面取得的成就，坚信在新的机遇面前，中兴通讯将在更大的舞台上展示实力，在全球企业中做最好的自己。

中兴通讯技术杂志社常务副总编黄新明介绍了2015年刊物发展情况，产学研办公室副主任李婷展示了产学研合作的最新成果，中兴通讯副总裁孙枕戈解读了“中兴通讯2016年技术白皮书”。针对这3个报告，各位来宾各抒己见，热烈讨论。

与会专家充分肯定了刊物及产学研工作所取得成绩，也希望刊物在科技期刊界发出更多声音，以促进更多刊物进步；同时，希望刊物参与国家期刊标准化制订工作。

一年一度的北京联谊会是个温暖的聚会，它拉近了刊物和业界专家的距离，成为产学研及杂志社对外交流的重要平台！

DOI:10.3969/j.issn.1009-6868.2016.01.012
网络出版地址:http://www.cnki.net/kcms/detail/34.1228.TN.20151202.1107.002.html

位置信息辅助的机间自组网路由协议研究

The Location Aided Routing Protocol in an Aircraft MANET

中图分类号:TN929.5 文献标志码:A 文章编号:1009-6868 (2016) 01-0046-004

史琰/SHI Yan
杨鹏/YANG Peng
(西安电子科技大学,陕西 西安 710071)
(Xidian University, Xi'an 710071, China)

摘要: 提出了一种适用于机间自组网的路由协议算法,该算法使用位置信息辅助计算节点间的链路持续时间,并以此链路持续时间作为拓扑稳定情况的预测。各节点则依据跳数最小原则和链路持续时间最长原则进行路由计算,并在条件允许的情况下,为网络中的节点建立两条路由。仿真结果表明,该算法能够满足机间自组网的高动态拓扑变化,提供良好的网络性能。

关键词: 自组织网络;链路持续时间;表驱动路由;多路径路由

Abstract: In this paper, an aircraft mobile ad hoc network (MANET) routing protocol is proposed. In this routing protocol, a new link parameter called link duration is computed according to the locations of aircraft and is used to forecast the stability of a wireless link. With the link duration, path duration can be obtained between any two nodes. A node computes each routing entry, satisfying both the least hops and longest path duration and finds another point disjoint path if possible. The simulation results show the routing protocol can adapt to network topology and improve network performance.

Key words: ad-hoc networks; link duration; table driven routing; multi-path routing

机间自组织网络是移动Ad Hoc网络(MANET)在航空通信领域的应用,其基本思想是:在一定范围内的飞行节点通过互相发送控制信息、感知信息等自动地建立起一个MANET[1]。在机间自组网中,飞行节点不但作为消息的收发节点,同时还在网络中担当路由器的功能,这使得机间自组网可以采用多跳的方式传输数据,扩大网络的覆盖范围。

机间自组网应用于民航通信可为空中交通管理提供新的技术[2],为航班提供通信保障[3];应用于军航通信可发挥抗毁、协同等优势,提升平台的战术效能[4]。与一般的自组织网络相比,机间自组网不但具有多跳、自组织、无中心等固有的特点,同时还具有节点分布场景广密度低[5]、网络拓扑的高动态性[6]、信道质量的不稳定性[7]、网络的异构性和临时性。

收稿日期:2015-11-20
网络出版时间:2015-12-02

1 位置信息辅助的机间自组网路由

1.1 机间自组网路由存在的问题

由于机间自组织网络具有节点快速移动、拓扑变化迅速的特点。在使用以往的基于最短路径的路由时,路由计算时只考虑了路径的长度。这在节点静止或节点低速移动的场景中能够适用,但是在机间自组织网络中,节点的快速移动会导致节点间的无线链路频繁通断。节点间链路的持续时间已成为影响路由的重要因素:距离最短的路径其链路持续时间可能很短,其在通信过程中的失效则会导致丢包率上升从而降低网络性能;链路持续时间长的路径可能增加路由的距离,加重网络中节点的负载,同时会增大信息传输的端到端时延。为了使所设计的路由协议适应节点的移动并能够使网络具有良好的性能,在路由算法中将采用最短路径原则和最长链路持续时间原则相结合。

1.2 位置信息辅助的链路持续时间计算

如图1所示,假设两节点A和B,其中节点A的经纬度分别为(ϕ_A, θ_A),速度为ν_A,航向为C_A,飞行高度为H_A;节点B的经纬度分别为(ϕ_B, θ_B),速度为ν_B,航向为C_B,飞行高度为H_B。以下关于角度的计算均是以正北方向为基准方向,节点B对于节点A的方位角为γ,两节点间的航向夹角为α,节点A的航向与两节点间连线的夹角为β,两节点A、B与地球球心形成的球心角为δ_{AB},

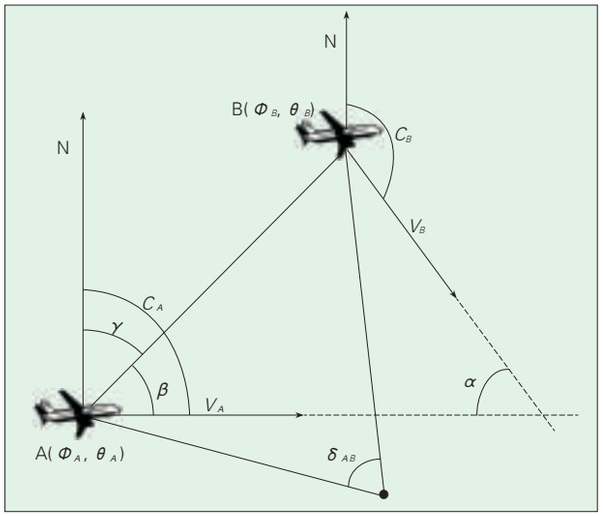


图1
节点间链路持续时间计算

节点间的最大通信距离为R。

根据两节点的经纬度信息，我们可以计算出两节点以地球球心为顶点形成的角距离：

$$\delta_{AB}=\cos^{-1}\left(\sin\theta_A\sin\theta_B+\cos\theta_A\cos\theta_B\cos\left(\phi_A-\phi_B\right)\right) \quad (1)$$

在计算两节点间的直线距离时，考虑到机间自组织网络中节点的飞行高度H（约10 km）相比地球半径L（6 400 km）要小很多，故在计算两节点间的直线距离时忽略了节点的高度（高度所带来的距离误差在1‰）。这样两节点A、B与地球球心近似形成一个等腰三角形，依照三角形中的边角关系可以计算出两节点间的直线距离：

$$S_{AB}=2\times L\times\sin\frac{\delta_{AB}}{2} \quad (2)$$

在计算出两节点A、B的直线距离后，再计算节点B对于节点A的方位角：

$$\gamma=\frac{\sin^{-1}\left(\cos\theta_B\times\sin\left(\phi_A-\phi_B\right)\right)}{\sin\left(\cos^{-1}\left(\sin\theta_A\times\sin\theta_B+\cos\theta_A\times\cos\theta_B\times\cos\left(\phi_A-\phi_B\right)\right)\right)} \quad (3)$$

节点A与节点B航向的夹角α：

$$\alpha=\left|C_A-C_B\right| \quad (4)$$

节点A的航向与两节点间连线的夹角β：

$$\beta=\left|C_A-\gamma\right| \quad (5)$$

在这里构造两个辅助变量$C(v_A,v_B,\alpha,\beta)$和$D(v_A,v_B,\alpha)$，其中：

$$C(v_A,v_B,\alpha,\beta)=2\times S\times\left[v_A\cos\beta-v_B\cos(\beta-\alpha)\right] \quad (6)$$

$$D(v_A,v_B,\alpha)=v_A^2+v_B^2-2v_Av_B\cos\alpha \quad (7)$$

最后，通过上述计算可以得到节点A、B间的链路持续时间：

（1）当时$v_A=v_B$且$\alpha=0$时，如果$S<R$，节点A和节点B在通信范围内，且保持相对静止，其链路持续时间可以认为是无穷；如果$S>R$，节点A和节点B无法通信，链路持续时间为0。

（2）当$v_A\neq v_B$或$\alpha\neq0$时，节点A和节点B间链路持续时间为：

$$T(v_A,v_B,\theta,\phi)=\frac{\sqrt{\left[C(v_A,v_B,\alpha,\beta)\right]^2-4(S-R^2)D(v_A,v_B,\alpha)}+C(v_A,v_B,\alpha,\beta)}{2D(v_A,v_B,\alpha)} \quad (8)$$

1.3 位置信息辅助的路由算法步骤

在本算法中，节点内部包含有2种类型结构表：网络拓扑表TE和节点路由表RT。

每个机间自组网节点在本地存储一张拓扑表TE，用于存储网络中各节点的位置信息。该拓扑表包含参数LINK_TIME以表明相邻节点之间的链路持续时间。如果节点i与节点j为邻节点，TE[i][j].IS_VALID_FLAG = 1可以表明两节点的邻居关系，并且TE[i][j].LINK_TIME可以表示节点i与节点j之间的链路持续时间。节点通过周期性地发送HELLO包的方式来进行拓扑表的建立和维护。HELLO包中会携带目前本节点已知的拓扑关系及链路信息。

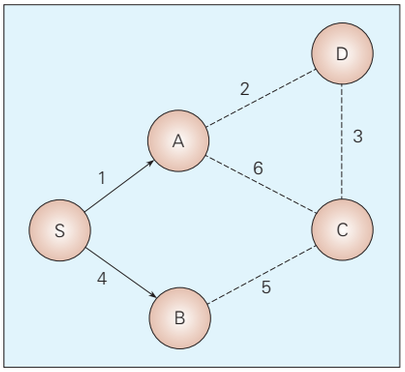
节点内部使用节点路由表RT记录到达其他节点的路由，并且在条件允许的情况下，为网络中的节点建立两条路由：RT[i][0]和RT[i][1]，其中i为目的节点（i =1,2,…,N, N为网络中节点数量），标志0为优先路由，标志1为备份路由。与网络拓扑表TE类似，RT[i][0].IS_VALID_FLAG =1表明该路由表项有效，RT[i][0].PATH_TIME记录该路由路径的持续时间。

机间自组网节点通过上文所述的方法获取网络拓扑信息并更新相应的拓扑表项，然后各节点根据本地存储的拓扑表来计算路由生成表驱动路由表。

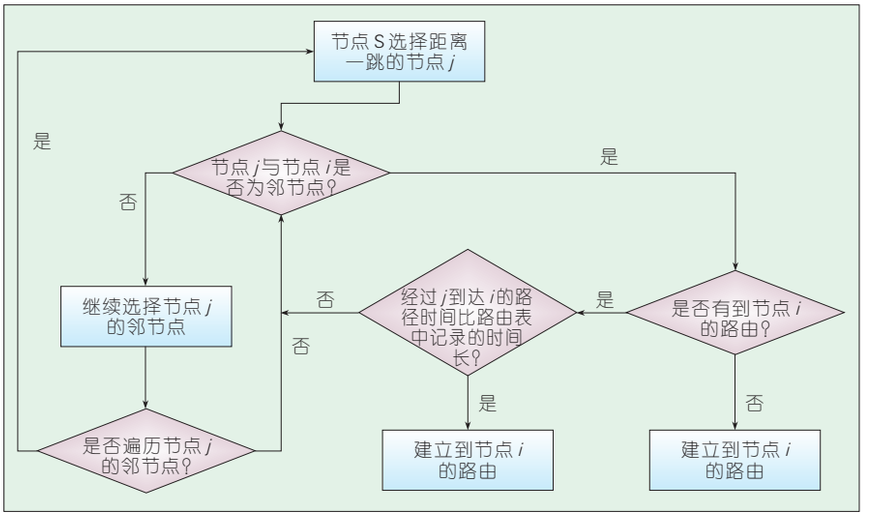
（1）假设本地节点为S，节点S首先初始化本地路由表RT[i] [j].IS_VALID_FLAG = 0（i =1,2,…,N, j = 0,1）；然后节点S再查找拓扑表TE内所有与自己为邻居的节点，即TE[S][j].IS_VALID_FLAG = 1，(j =1,2,…,N)，如果存在就更新节点j对应的路由表表项，并记录其与节点j之间的链路持续时间、距离、下一跳。如图2所示，节点S根据本地的拓扑表为相邻的节点生成路由表，图中节点S首先生成到节点A和节点B的路由。

（2）节点S根据各一跳节点的拓扑关系计算两跳范围内的路由，如图3所示。

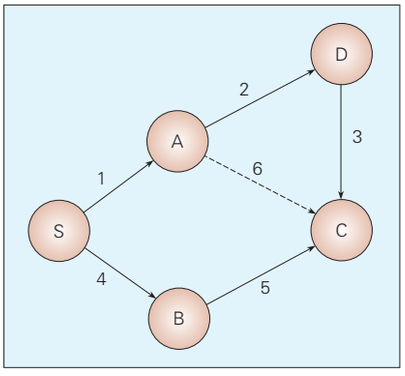
（3）节点S根据两跳节点的拓扑关系继续计算，并按照跳数的增加逐步扩散出去，直至到所有节点的路由都被计算出来。如图4所示，节点S



▲图2 节点表驱动路由的计算



▲图3 节点两跳范围内路由计算流程



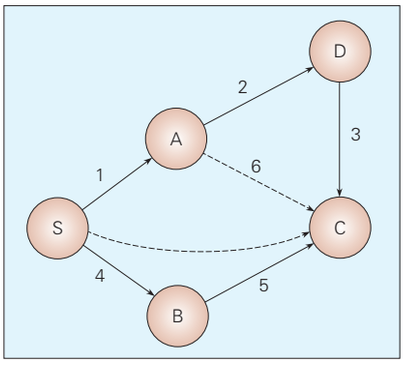
▲图4 节点表驱动路由的计算

根据本地的拓扑表中节点A和节点B的邻居关系生成两跳范围内的路由表，图中节点S通过节点A可以计算出两条跳数为两跳的路由，其中一条到节点C，另一条到节点D，同理节点S根据节点B的邻居关系计算到节点C的跳数为两跳的路由。由图中还可以看出节点S为节点C建立了两条路由，S到C的路由：优先路由为S→B→C，备份路由为S→A→D→C。

（4）在节点S计算路由表时，可能会出现节点S到某一节点j有多条路由，此时节点S依照图5所示原则对计算出的多条路由进行处理：

首先，节点S从计算出到节点j的多条路由中选择跳数最短的路由，当同时存在多条跳数最短的路由时，选择其中持续时间最长的路由作为优先路由，并将路由表项RT[j][0]按照上文所述方法更新；其次，当到节点j存在其他跳数次短但持续时间比优先路由时间长的路由时，将这条路由作为备份路由并更新路由表项RT[j][1]，同样当存在多条跳数次短的路由时，选择其中持续时间最长的路由作为备份路由。

如图6所示，假设图中的链路持续时间按其链路编号由大到小排列



▲图6 节点多路由选择策略示意

（链路1持续时间最长），节点S到节点A的链路持续时间最长且跳数最短，则节点S只为节点A建立一条优先路由：S→A。节点S到节点C的两条路由分别为：优先路由为S→B→C，备份路由为S→A→D→C。如图中所示，虽然链路S→A→C同样是两跳，但是由于其与链路S→B→C跳数相同且链路持续时间比S→B→C短，所以舍弃链路S→A→C。

总而言之，优先路由为节点S到节点j最短且持续时间最长的路由，备份路由为节点S到节点j次短但持

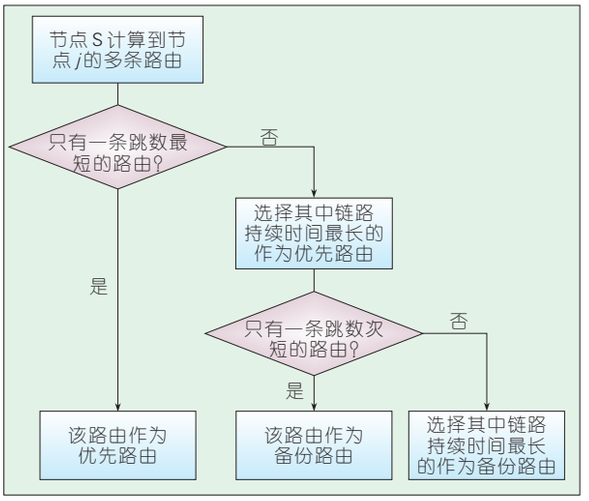


图5▶ 节点多路由选择策略

续时间比优先路由时间长的路由。当节点S使用优先路由与节点j进行通信时发现链路即将断开时，节点S切换备份路由进行通信，以此来保障节点间通信的连续性。

2 位置信息辅助的路由算法仿真结果

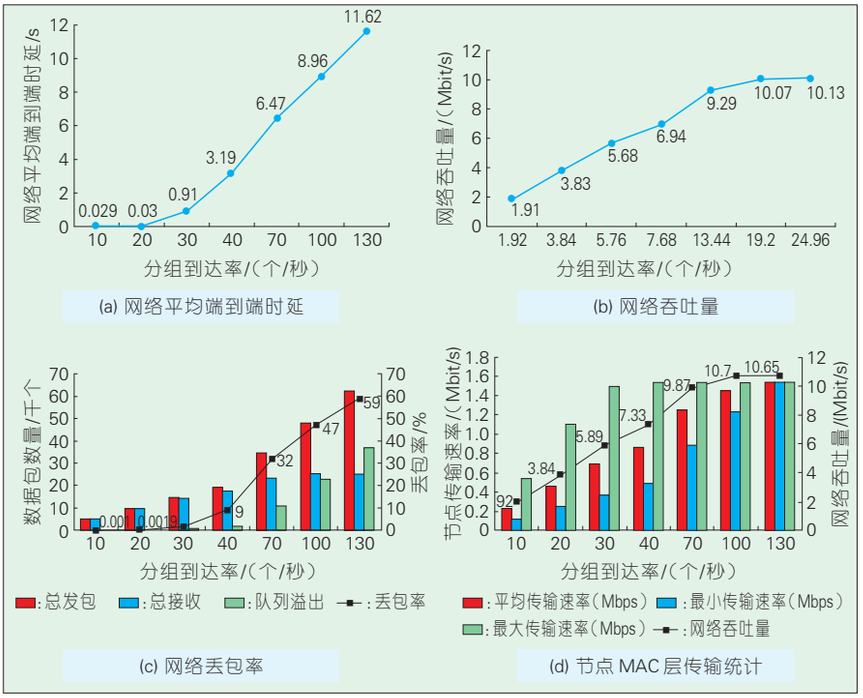
仿真软件使用OPNET14.5，其中各参数设置如表1所示。

▼表1 仿真参数设置

参数名	值
网络范围/ km	1 000 * 1 000
移动节点数目/个	16
节点最大通信距离/ km	200
仿真时间/ s	3 600
MAC层时隙大小/ms	2
数据分组长度/bit	1 200
收发信机信道速率/ (Mbit/s)	50

仿真中媒体接入控制层（MAC）采用时分多址（TDMA）的形式。时隙长度为2 ms，其中1 ms为发送数据，1 ms为保护间隔。时隙没有空间上的复用，节点发送周期为32 ms。物理层采用全向天线，信道速率为50 Mbit/s。节点每时隙内发送数据量上限为50 kbit，发送周期为32 ms。MAC层的发送速率上限为1.5625 Mbit/s。

图7仿真结果是在节点的移动速度固定为220 m/s，分组产生速率分别为10、20、30、40、70、100、130个/秒/节点，而MAC层缓存队列长度为1 000 pk下进行的。图7（a）为网络平均端到端时延，随着网络负载的增加，分组平均端到端时延由0.03 s增加到11.62 s；图7（b）为网络吞吐量与网络负载之间的关系，图7（c）为网络丢包率与网络负载的关系，由这两幅图可以看出随着负载的增加网络吞吐量逐渐增加并趋于稳定在10 Mbit/s，而网络丢包率增加到59%；由图7（d）为节点MAC层的传输能力统计曲线，可以看出网络中每个节点的MAC层均达到了其传输能力的上限，因此限制了网络性能的提升。



▲图7 位置信息辅助的路由算法性能仿真

3 结束语

机间自组织网络具有节点快速移动、拓扑变化迅速的特点，机间自组织网络中的路由很大程度上受到这些特点的影响。位置信息辅助的最短路径原则和最长链路持续时间原则相结合的路由算法，可以降低高动态变化的网络拓扑对路由的影响。在网络拓扑允许的情况下，通过使用优先路由和备份路由的方法，保障了数据信息在节点间传输时不受链路通断的影响。

参考文献
[1] EHSSN S，ABBAS J. The Global in-Flight Internet [J]. IEEE Journal on Selected Areas in Communications, 2006, 24(9): 1748-1757
[2] MAGGIE X C. Connectivity of Ad Hoc Networks for Advanced Air Traffic Management [J]. Journal of Aerospace Computing, Information and Communication, 2004, 1(5): 225-238
[3] HU D T, SHIGERU S. A Proposal of Relaying Data in Aeronautical Communication for Oceanic Flight Routes Employing Mobile Ad Hoc Network[C]// 2009 First Asian Conference on Intelligent Information and Database Systems, Washington DC, USA, 2009
[4] 韩勇，陈强，王建新.机载网络技术综述[J].电讯技术，2008,48(8):111-114
[5] YANG W. Fundamental Issues in Systematic Design of Airborne Networks for Aviation[C]// IEEE Aerospace Conference, 2006
[6] JUSTIN P R, ABDUL J, EGEMEN K C, et al. High-Dynamic Cross-Layered Aeronautical Network Architecture [J]. Aerospace & Electronic Systems IEEE Transactions on, 2011, 47(4):2742-2765
[7] ERIK H. Aeronautical channel modeling [J]. IEEE Transactions on Vehicular Technology, 2002, 51(2): 254-264. doi: 10.1109/25.994803

作者简介

史琰，西安电子科技大学副教授，综合业务网理论及关键技术国家重点实验室专职研究人员；研究方向为认知网络、无线分布式组网与技术；近五年获省部级奖2项；授权发明专利10余项，发表学术论文20余篇。

杨鹏，西安电子科技大学硕士生，研究方向为无线自组织网络，申请发明专利1项。

DOI:10.3969/j.issn.1009-6868.2016.01.013
网络出版地址:http://www.cnki.net/kcms/detail/34.1228.TN.20151127.1621.002.html

无线核心网的TCO分析方法研究

TCO Analysis for Wireless Core Network

中图分类号:TN929.5 文献标志码:A 文章编号:1009-6868 (2016) 01-0050-004

摘要: 基于通信行业的总体拥有成本(TCO)模型和核心网的特殊性,提出了无线核心网专有的TCO模型。该模型的特殊性包括:互联互通测试费用的复杂性、培训费用的必要性。通过引入新供应商的决策场景,采用成本项纳入工作流的分析方法,提出以人力投入成本和时间为主要度量手段的TCO定量度量方法。通过新老供应商并存场景的分析,指出业务部署和运维会对TCO产生复杂的影响。

关键词: 无线核心网; TCO; 成本收益预测; 成本管理

Abstract: Based on the model of total cost of ownership (TCO) in the communication industry and the special nature of core networks, the TCO model of wireless core network is put forward. The drawbacks of the model include complex interconnection testing and training cost. By introducing the decision scene of new supplier's, and adopting the analysis method in which cost items are included in the workflow, the TCO quantitative measurement method based on human input cost and time as the main measure method is put forward. Through analysis on the scene which includes the new and old vendors, it is pointed out that business deployment and operation and maintenance have a complex effect on TCO.

Key words: wireless core network; TCO; revenue-cost estimation; cost management

史庭祥/ SHI Tingxiang
田会芹/ TIAN Huiqin
(中兴通讯股份有限公司 南京研究所,江苏 南京 210016)
(Nanjing R&D Center, ZTE Corporation, Nanjing 210012, China)

无线网络的供应商选择,一直是困扰运营商的难题,这其中往往涉及技术和商务的多轮比较,评估决策时间超出1年也很常见,其原因在于:运营商不仅需要比较产品质量和价格,更需要综合考察供应商提供设备和服务的总体拥有成本(TCO)。

基于TCO的研究有助于运营全周期的成本控制,然而业界对于无线核心网的TCO分析并不多见,文章正是对该方面做出一些研究和探索。

1 通信行业的TCO模型

随着通信市场成熟度的逐步提高,移动数据承载能力也随之快速提升,IT互联网企业(如SKYPE、QQ、微信、Twitter、WhatsApp等)从以往与运营商合作的关系变成竞争者的关系,这使得运营商既面对运营商之间的明处竞争,又无法逃避和IT企业之间的暗处较量。在业务收入(如语音和短信)被IT分流的情况下,运营商不得不更加严格地控制设备和服务的采购成本。与此同时,全球运营商纷纷采用设备集中招标,如中国的三大运营商从2008年开始的3G招标中几乎每个项目都是集团统一招标和采购,而海外运营商则更多采取兼并换股等方式,集中度大幅提升,实现集约化经营、集中化采购管理,这些调整都有利于降低采购成本。然而在降低采购成本策略实行的初期,人们往往只关注低的设备价格,这可能导致总购置成本升高。比如,供应商表面上降低设备价格等当期成本而赢得订单,却为弥补"损失"提高软件和服务等远期费用,隐藏客户投入的显性和隐形成本。

为此,探讨设备之外降低成本的方法对运营商选择供应商也非常重要[1]。TCO作为一种全新的理念[2],旨在让人们更好地理解与供应商发生商品和服务交易的真实成本,该真实成本被用于采购管理,甚至开始应用于基于TCO的供应商选择[3]。近些年,其他国家的一些高端运营商已普遍开始使用TCO的方式,全方位地进行采购成本分析,以便从设备的整个生命周期选择最为经济的设备以及方案。

对于TCO的分析,我们不仅要考虑网络设备的购买成本(CAPEX),还要考虑一定年限内的运营成本(OPEX)。对于通信行业来说,CAPEX包括设备购买成本和设备部署成本,例如设备软硬件的购买成本、运营保障成本、搬运成本、安装调测成本、工程土建成本等;OPEX包括设备运营成本和设备运维成本,例如站点租金成本、设备电费成本、设备

收稿日期:2015-10-20
网络出版时间:2015-11-27

维修成本、人力成本等。有时可能还会计算机会成本，例如设备故障所来带的营业损失、客户满意度损失[4]等。通信行业的TCO模型，一般如图1所示。

2 无线核心网的TCO模型

如何利用通信行业的TCO模型和工具，为运营商的投资决策提供有力依据，已得到主流设备供应商的重视，并在无线接入网中有所实践[5]，但在核心网方面的深入研究还比较少。

基于多年项目经验，我们认为：相比于设备成本，我们更需要关注多家无线核心网的运维和替换成本的比较。作为通信系统的子系统之一的无线核心网，它的TCO模型既要服从通信行业的通用模型，又要满足核心网的一些特殊性，其TCO模型如图2所示。

无线核心网TCO模型的特殊性对TCO的研究至关重要。无线核心网TCO模型的特殊性包括：互联互通测试费用的复杂性、培训费用的必不可少。

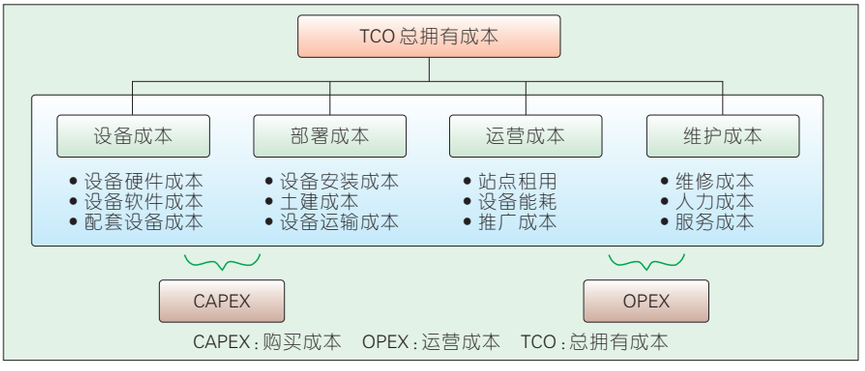
（1）互联互通测试费用的复杂性

互联互通测试（IOT）的复杂性是核心网区别于其他网络子系统的重要特征。由于核心网在网络中起到枢纽的作用，它几乎和所有子系统都有互通关系，这使得IOT费用存在一定程度的难以预料。IOT费用既包括设备商投入的直接费用，又包括运营商投入的直接费用以及业务部署、维护和影响项目工期的隐形成本。根据互联互通对象的不同，IOT可以分为两种：

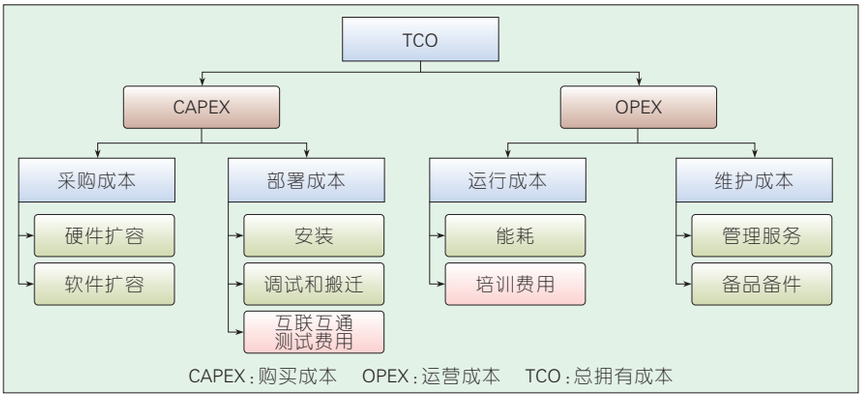
- 核心网和通信网络的其他子系统的IOT，如无线接入子系统、业务子系统等。
- 核心网之间的IOT。

（2）培训费用必要性

核心网属于高技术复杂度的产品，运营费用中的培训费变得必不可少，甚至不可低估，这是区别于一般技术产品的业务运营特点。该费用高低决定于项目方案的复杂程度，以及现有维护人员的技术水平。项目方案的复杂程度由网元数量、功能点规模等决定；维护人员技术水平一般和新产品、新方案的应用相关，在一些情况下我们认为引入新厂家将很大程度上增加该费用，并带来业务运营的进度风险。



▲图1 通信行业的TCO模型



▲图2 无线核心网的TCO模型

3 无线核心网的TCO度量

核心网属于软硬结合，且价值偏重软件的产品，针对每个成本项，我们可以从人力投入的角度来分析其成本，包括如下几方面要素：

（1）某项成本的规模，如网元数、功能点数、接口数等。

（2）投入资源的规模，如工程师数、工作天数、占地面积。

（3）人力成本，如单位时间的人员薪资。

（4）资源成本，如设备功耗、单位时间的电费、机房租金等。

（5）调节系数，如风险系数等。

为了更清晰地说明如何度量TCO，我们以引入新供应商的场景做分析，如图3所示。引入新供应商无疑会带来更高的成本和风险，因此首先需要细化工作流，该工作流包括如下7个步骤，其中每个步骤可看作是一个或多个成本项：

（1）测试床和IOT。

（2）新设备安装、调试和集成（ICI）。

（3）新设备试验局运行。

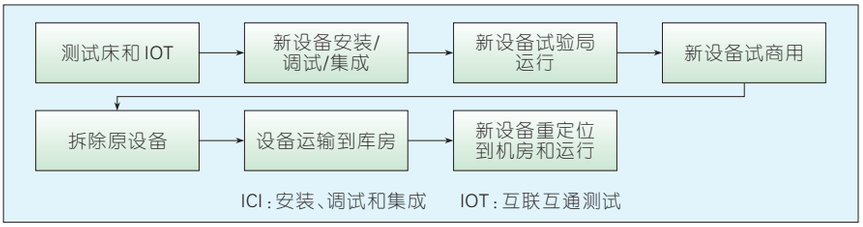
（4）新设备试商用。

（5）拆除原设备。

（6）设备运输到库房。

（7）新设备重定位到机房和运行。

接下来，我们需要对工作流的每个步骤给出定量分析，主要从成本和时间两方面进行考虑。为方便理解，我们主要分3部分描述，第1部分说明工作流的第1个步骤，即测试床和IOT；第2部分描述工作流的剩余6个



▲图3 引入新供应商的工作流

步骤；第3部分说明在该工作流完成后，从新设备的业务部署和运维角度出发，该如何进行TCO分析。

（1）工作流的第1个步骤：测试床和IOT

在新设备引入期间，运营商需要考虑和新供应商的合作，并安排相关工程师予以配合（假设1个网元有1位工程师），测试床和IOT的成本计算方法如下。

- 人力成本：网元数（IOT对应接口数）×配合天数×每工程师的日薪。
- 资源成本：占用机房的每天租金和维护费用×配合天数+设备每天功耗×配合天数。
- 时间成本：即配合天数。

最终给出的成本还需要乘以风险系数，测试床一般取1～2，IOT一般取4（IOT的风险远大测试床）。IOT涉及的接口情况，如表1所示。

（2）工作流的剩余6个步骤

一般来说，新设备的安装、调试和集成大概需要1个月，考察合格后新设备试验局运行的观察期为1～3个月，正式商用后的观察期为1个月，然后开始老设备的替换搬迁，之后再进行拆除和入库。从新设备在新站点上的设备安装、调试和集成到老设备移交库房的工程周期估算如图4所示。依据项目经验，对于单站点单设备，从设备安装、调试和集成到老设备移交库房的工程周期需要51天；若考虑到多网元多站点的并行，由于多网元和多站点相互配合需要时间，总的工程周期大约为：1月+3月+1月+51天/站点=200天。

新设备启用的成本估算方法为：人力或资源成本×配合天数。

（3）业务部署和运维的TCO分析

引入新供应商，会带来业务部署和运维上的差异，同时由于新老供应商的共存情况，业务部署和运维的复杂程度和TCO可能远大于估算，我们将从两个方面进行分析。

（a）业务部署TCO分析

新的供应商和现网供应商共存或暂时共存的阶段，必然带来双倍的部署工作量。另外，两个供应商在业务部署中存在协同问题，如图5所示，红色部分是引入新供应商后增加的流程。两个供应商的业务协同工作由运营商负责，当两个供应商的业务协同不成功的时候流程需要回溯直到成功为止，这个过程需要增加开发和确认的成本，以及进度延误等隐性成本。

一般来说，中小规模的业务部署的准备周期，包括需求获取、需求评估、需求开发以及需求确认4个环节。而多个供应商共存情况，则增加了“多供应商之间的业务协同”环节，以及可能的需求二次开发，使得总周期大约增加1/3甚至1倍。若单个供应商的中小规模的业务部署准备周期按3个月计算，两个供应商的业务部署周期则需要4～6个月。

▼表1 IOT接口情况示例

序号	互联互通测试项	序号	互联互通测试项
1	MSC A接口测试用例	11	STP管理禁止
2	MSC Nc/BICC接口测试用例	12	STP ASP状态维护
3	MSC CAP接口测试用例	13	OCS接口测试用例
4	MSC Nc/ISUP接口测试用例	14	VAS接口测试用例
5	MSC MAP接口测试用例	15	CC&B接口测试用例
6	HLR MAP接口测试用例	16	NGCC接口测试用例
7	PS Gs接口测试用例	17	NOC接口测试用例
8	PS Gb接口测试用例	18	SMS接口测试用例
9	PS Gc接口测试用例	19	VMS接口测试用例
10	STP 7号信令管理	20	IT管理和业务开通接口

ASP：应用服务器(AS)进程
BICCv：承载无关的呼叫控制协议
CAP：智能业务(CAMEL)应用部分
CC&B：客户服务与计费系统
ISUP：综合业务数字网(ISDN)用户部分
IT：信息技术
MAP：移动应用部分
MSC：移动交换中心
NGCC：下一代呼叫中心
NOC：网络维护中心
OCS：在线计费系统
PS：数据域
SMS：短消息系统
STP：信令转接点
VAS：增值服务
VMS：语音信箱系统

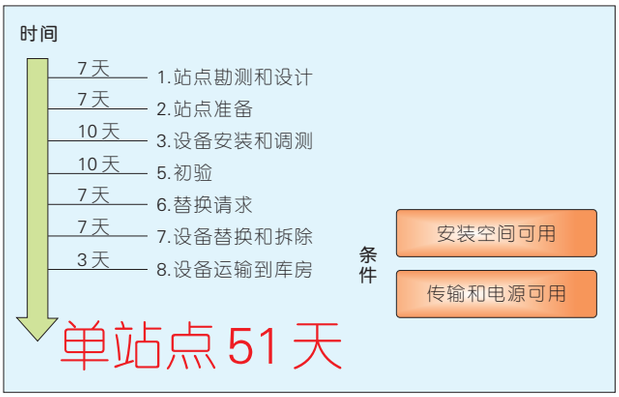
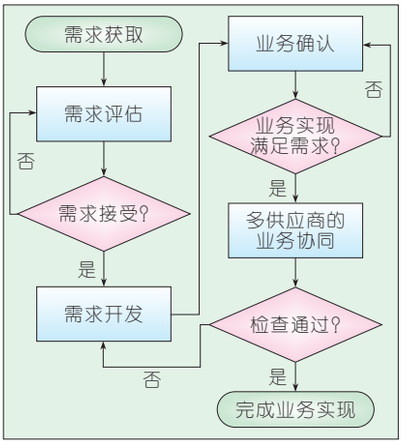


图4▶ 从新设备安装、调试和集成到设备运输到库房的时间估算



▲图5 多供应商的业务部署准备流程

业务部署的总体资源和时间消耗计算如下：

• 运营商的人力资源。单一供应商的业务部署需要工程师3人/月，两家供应商则需要8～12人/月。

• 业务部署周期。单一供应商的业务部署时间需要1周，两家供应商则需要2周。

• 业务部署对应的人力资源。单一供应商的业务部署时间需要1人/周，则两家供应商需要4人/周

（b）运维TCO分析

运维TCO分析包括运行和维护两部分，而运行部分又包括设备功耗和培训费用。

• 设备功耗和设备本身相关，一般由厂家提供设备的功耗标准数据，以及本次项目准备投入的设备的功耗表。除当期功耗分析外，还需要预测3～5年后容量扩展带来的功耗变动，以及下一个技术演进周期的功耗预测。

• 引入新供应商会导致培训费用显著增加，相比于原供应商的系统更新带来的培训需求，相关培训费用预算更大。特别说明的是，即使供应商在竞争环境下采用激进的商务策略，给予超低的培训报价，仍然无法清晰地测算出运营商为此付出的相关人员的劳务费用。

• 项目成功要素中的进度因素，维护人员和新产品新方案的磨合带来的其他风险，这些因素和风险往往难以测算。即使采用运维托管模式，上述风险通过收益分享方式转移给供应商，并通过供应商的主动运维提升效率，上述风险也并未能消除[6]。由此可知：引入新供应商，培训费用上升是大概率事件。

运维阶段的TCO分析包括专业管理服务和备板备件两部分。

• 专业管理服务成本。依据设备特点，专业管理服务的TCO分析主要包括故障定位服务、软件升级服务等。和业务部署类似，在同样的服务范围下，运营商要求新供应商以更高的运维效率提供更低廉的专业服务。但在两家供应商并存的情况下，一般来说，不同于分区域建设原则下无线厂家之间的协同成本偏低的特点，核心网往往有难以预测的协同成本，至少在搬迁过程或两家供应商并存期间，两家核心网的专业服务的总成本往往高于单家核心网，相关分析可参照“业务部署”中的相关部分。

• 备板备件成本。备板备件成本一方面和设备有较大相关性，如设备间不能共享板件将增加当期以及未来备板备件的采购成本；另一方面，若设备的平滑演进方案不佳，同样会增加未来演进到新设备时的备板备件成本，因此这部分成本分析，需要供应商提供当期和演进各阶段的备板备件产品配置和报价情况。另外，和单供应商相比，多供应商场景中，每个供应商承接的网络规模变小，即单一供应商承接的大网络被拆分为多个小网络，由不同供应商交付，因而备板备件成本一般会上升，而且一旦某个供应商退出网络，其备板备件成本将成为运营商的沉没成本。然而，供应商策略在备板备件成本分析上并非总是占劣势，比如供应商选择存在区域化需求时（包括服务和运营的区域化），备板备件则也需要匹配这个要求，即只向单供应商采购备板备件。这种情况下，供应商策略并不会因为设备共享带来更低的成本，反倒是因为供应商数量不足而失去议价优势。

4 结束语

近年来，TCO分析已经成为供应商必备的决策工具。在文章中，我们基于无线核心网的项目实践，将客户需求融于TCO分析，提出了核心网TCO模型和度量方法等，并对新设备商引入场景下的TCO决策进行重点探讨。希望能为运营商和设备商的核心网从业人员提供TCO分析方法，以便在核心网设备选择时有更多的参考依据。

除在引入新供应商的场景深化定量分析外，我们还可以针对新建网络、网络替换改造等其他场景作TCO分析，这些都可作为继续研究的课题方向。

参考文献

[1] 艾默生. 降低TCO：网络能源厂商竞争大趋势[J]. 电信网技术, 2009(8): 86–87
[2] LM ELLRAM, SP SIFERD. Purchasing: The Cornerstone of the Total Cost of Ownership Concept [J]. Journal of Business Logistics, 1993, 14(1), 163–184
[3] 李文霞, 邸颖, 佟晓莉. 采购中的总体拥有成本战略[J].辽宁经济, 2006(5): 110–112
[4] 潘丽. 电信设备采购的TCO管理研究 [D]. 北京邮电大学, 2008
[5] 利文. 爱立信TCO方案助运营商降低网络成本最新企业一体化通信平台上市[J]. 通信企业管理, 2006(8): 60–60
[6] ETISALAT. 运维托管实现快速超越[EB/OL]. http://www.c114.net/topic/2644/a580228.html

作者简介

史庭祥，中兴通讯股份有限公司高级工程师；从事核心网市场规划和管理工作，主要研究方向为无线、核心网、虚拟运营及其关键技术，以及相关市场策划理论和实践；发表论文3篇，已获授权发明专利10余项，国际专利2项。

田会芹，中兴通讯股份有限公司高级工程师；从事业务软件的产品规划工作，主要研究方向为业务软件、融合通信及其关键技术；已获授权发明专利3项以上。

DOI:10.3969/j.issn.1009-6868.2016.01.014
网络出版地址:http://www.cnki.net/kcms/detail/34.1228.TN.20151224.1649.004.html

基于号码携带的VoLTE网络互通研究

VoLTE Network Interworking Based on Number Portability

中图分类号:TN929.5 文献标志码:A 文章编号:1009-6868 (2016) 01-0054-005

缪永生/MIAO Yongsheng[1]
倪明/NI Ming[2]
(1.中兴软件有限责任公司,江苏 南京 210000;
2.中兴通讯股份有限公司,广东 深圳 518057)
(1.Nanjing ZTE Software Co., Ltd., Nanjing 210000, China;
2.ZTE Corporation, Shenzhen 518057, China)

摘要: 认为VoLTE技术能够提供高清语音、视频等多媒体业务,是LTE下唯一的、端到端的目标语音解决方案。针对号码携带的国家/地区中各运营商VoLTE网络之间互通,提出了4种方案以及部署建议,从而实现了不同运营商VoLTE网络之间互通,对全球VoLTE网络互通有较强的借鉴和指导意义。

关键词: VoLTE;IMS;互通;号码携带

Abstract: Voice over LTE (VoLTE) technology is used for multimedia services such as high definition (HD) audio and video calling. VoLTE is the only target solution for end-to-end audio service based on LTE. For multi-VoLTEs interworking for supporting number portability service's in the country/region, we give four effective solutions for different scenes and stage, as well as its deployment suggestion according to the solutions. It is helpful for different operator VoLTE network interworking.

Key words: VoLTE; IP multimedia subsystem (IMS); VoLTE interworking; number portability

演进到第4代无线数据网络对运营商来说是大势所趋[1],基于LTE的语音业务(VoLTE)能够提供高清语音业务与消息、高清视频等多媒体业务,以及同互联网业务进行融合,符合电信网络演进方向。全球移动通信系统协会(GSMA)和下一代移动通信网络联盟(NGMN)都已经宣布将VoLTE作为业界长期演进/(LTE)下唯一的、端到端的目标语音解决方案[2]。而目前只有同一个运营商的客户才能够使用VoLTE业务,为了应对过顶传球(OTT)的竞争,给客户更好的业务体验,不同电信运营商之间的VoLTE互通的需求尤为紧迫。

文章在深入研究VoLTE网络架构和流程基础上,重点研究了支持号码携带业务的国家和地区的不同运营商VoLTE网络之间的互通要求,提出了一些解决方案和部署建议。

收稿日期:2015-11-15
网络出版时间:2015-12-24

1 VoLTE网络互通概况

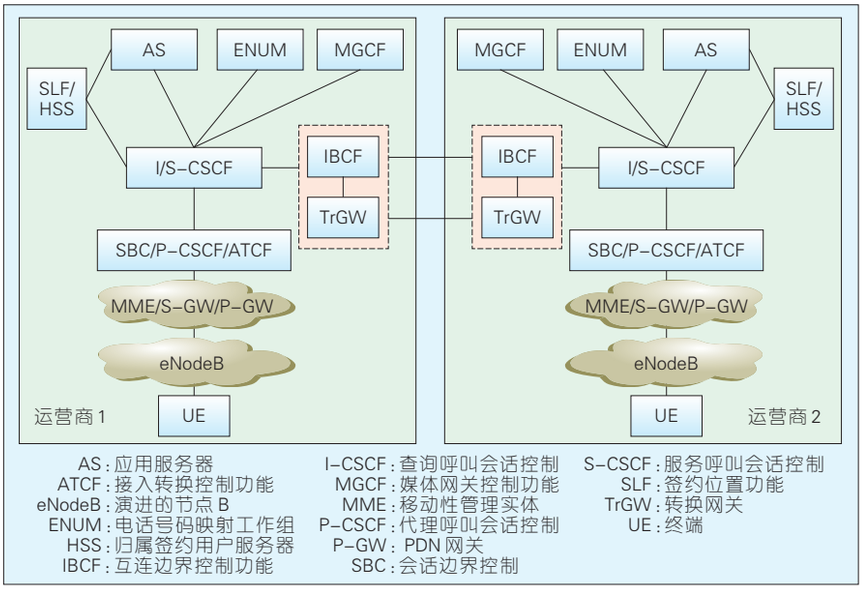
1.1 VoLTE系统架构

在VoLTE系统架构中,由LTE和演进的分组网(EPC)系统提供承载,由IP多媒体子系统(IMS)系统提供业务控制,实现端到端的基于分组域的语音、视频业务。IMS作为控制核心,可以提供和电路域类似的语音业务及其补充业务,包括号码显示、呼叫转移、呼叫等待、会议电话等。VoLTE网络典型系统架构如图1所示。图1中主要包括终端(UE)、LTE(eNodeB)、EPC(移动性管理实体(MME)、服务网关(S-GW)、PDN网关(P-GW))、IMS(接入会话边界控制(A-SBC)、代理呼叫会话控制(P-CSCF)、接入转换控制功能(ATCF)、查询呼叫会话控制(I-CSCF)、服务呼叫会话控制(S-CSCF)、应用服务器(AS)、电话号码映射工作组(ENUM)、媒体网关控制功能(MGCF)、签约位置功能/归属签约用户服务器(SLF/HSS))等网元。其中呼叫会话控制(CSCF)实现了呼叫控制,AS实现了类似2G/3G的多媒体电话补充业务[3-5]。

1.2 VoLTE与VoLTE网络互通

如果主、被叫用户均采用VoLTE业务,LTE和EPC网络对于VoLTE业务来说相当于接入网,并不参与网络互通。真正实现运营商间的VoLTE网络互通的是IMS系统,也可以说VoLTE网络的互通要求其实就是IMS网间的互通要求[6]。

通常对于不支持号码携带的国家/地区,跨运营商的VoLTE互通,能够根据被叫号码的信息,获知被叫用户所归属的运营商,从而实现不同运营商之间的互通。目前不同运营商VoLTE网络之间互通有如下的几种方案:



▲图1 VoLTE系统架构

（1）主叫IMS网络完成主叫业务后，主叫IMS网络的S-CSCF，根据被叫号码的号段信息，路由业务请求到被叫号码归属的IMS网络[7-10]。

（2）出于安全考虑，跨运营商之间都会部署边界网关。主叫IMS网络完成主叫业务后，主叫IMS网络的互连边界控制功能（IBCF）根据被叫号码的号段信息，选择被叫号码归属的IMS网络，并将业务请求路由到被叫归属的IBCF网元。其中，信令面的互通网元是IBCF，互通协议采用会话发起协议（SIP）；媒体面的互通网元是转换网关（TrGW），互通协议采用实时传输协议（RTP）。

（3）由于电路交换域（CS）网络是各移动运营商成熟的网络，VoLTE与VoLTE网络可以通过主叫IMS网络的S-CSCF将业务请求路由到MGCF，由主叫电路域网络路由到被叫电路域网络，再由被叫电路域网络路由到被叫IMS网络。这种方案的好处是通过电路域来路由保护了已有的投资，缺点是无法实现跨运营商的VoLTE高清语音、视频业务。

方案1和方案2类似，可以实现不支持号码携带的国家/地区不同运营商VoLTE网络之间互通，方案3缺点是无法实现跨运营商的VoLTE高清语音、视频业务。

1.3 号码携带下VoLTE网络之间互通

在网络支持号码携带业务后，用户的号码和运营商不再具有简单对应关系，现有网间互通根据被叫用户号码选路的方式被颠覆。号码携带可以是携入或者携出，号码携带业务开展后，对语音业务路由的判断采用被叫号码+路由号码（RN）综合判断的方式进行，对于签约号码携带业务的被叫用户则代之以RN为依据进行后续接续路由。对移动网络需要由主叫网络通过与号码携带数据库（NPDB）交互信息来判断用户是否签约号码携带业务以及RN信息，以便主叫网络根据RN信息选择正确的归属运营商网络。移动网络通常部署移动号码携带（MNP），MNP中保存了号码以及携入携出和RN等信息。

支持号码携带的国家和地区，跨运营商VoLTE网络之间的互通，就无法简单通过被叫号码进行路由。虽然电路域是成熟网络，支持查询MNP并针对被叫号码+RN完成路由，但是如果VoLTE网络之间通过传统电路域互通，由于承载的限制，则无法实现高清语音、视频业务。

在移动互联网时代，面对OTT的市场竞争，很多移动运营商正在加紧部署VoLTE网络，不同移动运营商也在加强合作，2014年AT&T和Verizon共同宣布2015年正式实现VoLTE跨网兼容和互通。其他国家和地区如香港CSL、HKT、数码通等运营商之间也在进行VoLTE互通测试或商用。因此号码携带的国家/地区各运营商VoLTE网络之间互通方案显得尤为重要和紧迫，必须能够实现支持号码携带的不同运营商VoLTE网络之间互通，从而实现VoLTE高清语音、视频业务。

2 VoLTE互通方案分析

2.1 两个VoLTE网络间互通方案

2.1.1 方案1

对于支持号码携带的国家和地区，如果不同的运营商之间进行VoLTE网络之间互通，那么就需要VoLTE和VoLTE直接互通，以保证高清语音、视频业务的正常开展。如果仅有两个运营商之间VoLTE互通，则可以采用下面提供的ENUM互查解决方案。假设运营商1的VoLTE用户呼叫运营商2的VoLTE用户，信令和媒体流向可参考图2。

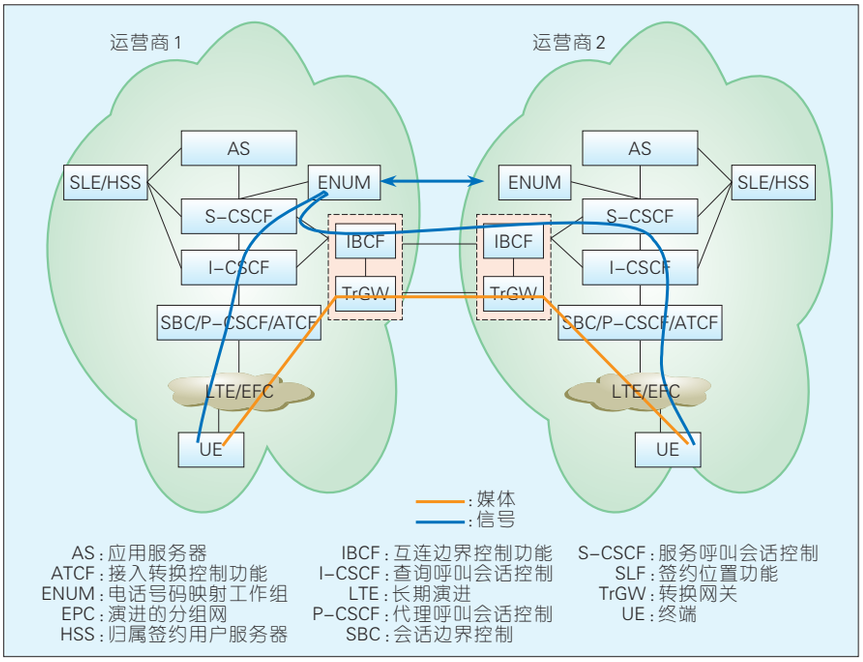
（1）运营商1的主叫IMS网络完成主叫业务后，主叫S-CSCF查询主叫网络的ENUM。

（2）运营商1的ENUM在本地查找被叫用户，如果查询不到则向运营商2的ENUM查询。

（3）运营商2的ENUM返回查询结果给运营商1的ENUM，然后运营商1的ENUM将查询结果返回给S-CSCF。

（4）主叫IMS网络的S-CSCF/IBCF根据ENUM返回的结果，将业务请求路由到运营商2的IMS网络，实现两个运营商之间的VoLTE互通。

由于支持号码携带业务，号码无



▲图2 ENUM互查方案

法区分不同的运营商，运营商1的ENUM在本地查询后，在对方开放ENUM查询接口的前提下，再去尝试查询运营商2的ENUM，这样就能确认被叫是否是运营商2的VoLTE用户，实现两个VOLTE网络的互通。

2.1.2 方案2

方案1需要ENUM在本地查找无果后，再去对端IMS网络查询ENUN。在方案1的基础上还有一种类似的解决方案2，需要S-CSCF定制处理逻辑，即S-CSCF首先查询本地ENUM，如果查询不到则再查询运营商2的ENUM，S-CSCF根据运营商2的ENUM返回结果，将业务请求路由到运营商2的IMS网络，能达到方案1同样的效果。方案2的如图3所示。这两个方案都需要运营商给对方IMS网络开放ENUM查询接口。

2.2 多个VoLTE网络间互通方案

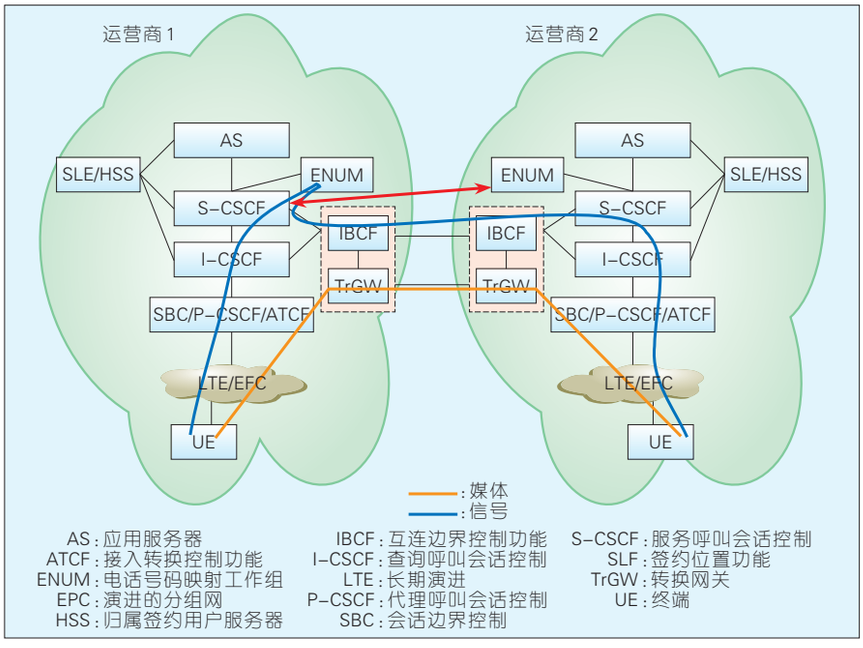
2.2.1 方案1

通常一个国家/地区，移动运营商会多于两个，不同运营商之间的VoLTE互通会存在更多的困难。在支持号码携带的国家/地区，如果借鉴两个VoLTE互通方案，本地的ENUM需要不断地尝试其他运营商的ENUM，但效率低、时延大，方案不可行。

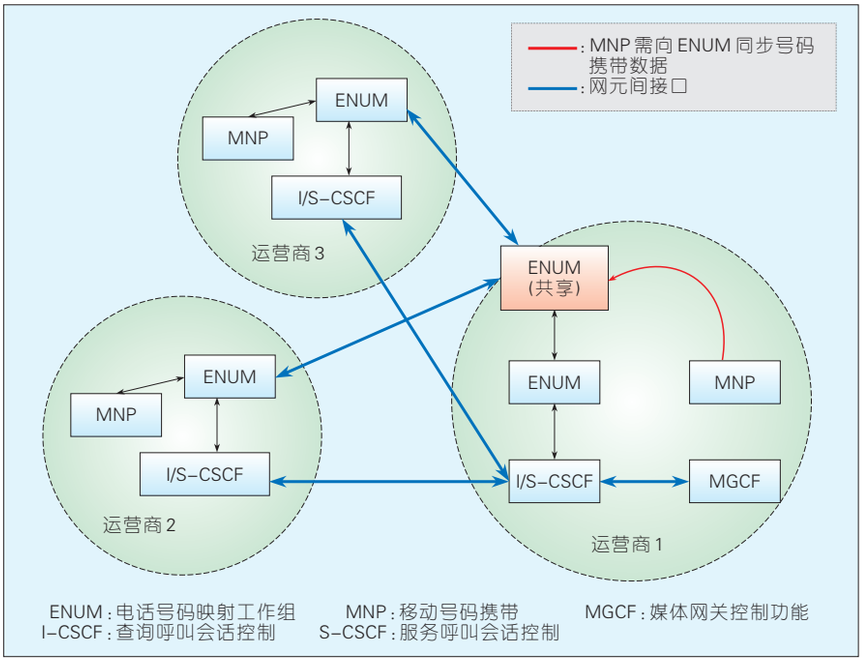
另外，传统的电路域移动网络，为了解决号码携带下的互通，通常必须部署移动号码携带（MNP），电路域网元查询MNP得到号码和RN信息，以便路由到正确的归属运营商。IMS网络目前和MNP网元没有直接的接口，如果IMS网络定制MAP接口直接查询MNP，一方面会增加IMS的复杂度；另一方面则会增加现网MNP的负荷。

假设电路域后续不再查询MNP，直接由IMS域查询MNP，但是这样对现网改动较大，且不利于保护运营商已有投资。因此在VoLTE网络建设初期阶段，我们提出了一种新的解决方案，实现了支持号码携带的国家/地区的多VoLTE网络之间的互通。

如图4所示，新方案需要部署一套共享ENUM，该逻辑网元可以和IMS中ENUM网元进行合设或者分设。共享ENUM从MNP中通过文件传输协议（FTP）同步文件方式，得到号码/号段归属运营商RN信息。当运营商1的主叫IMS网络完成主叫业务后，主叫S-CSCF查询主叫网络的ENUM。如果本地ENUM无查询的用户信息，则本地ENUM进一步查询共享ENUM，共享ENUM根据从MNP得



▲图3 S-CSCF查询方案



▲图4 共享ENUM互通方案

到的用户归属运营商RN信息，决定去相应运营商的ENUM查询。假设被叫号码现在归属运营商3，那么共享ENUM直接查询运营商3的ENUM，并返回相应的查询结果，主叫IMS网络通过查询结果将业务请求路由到运营商3。通过上述架构，实现了多VoLTE网络的互通，保证了高清语音、视频业务的正常开展。具体流程说明如下：

（1）MNP定时向共享ENUM同步FTP数据文件。

（2）共享ENUM根据从MNP同步的数据，得到号码/号段对应的运营商信息，即根据同步数据中RN信息获知归属运营商信息。

（3）主叫IMS网络S-CSCF针对被叫号码，查询本地ENUM，ENUM在本地查找无果后，进一步查找共享ENUM。

（4）共享ENUM根据从MNP得到的号码归属运营商信息，查询用户归属运营商开放的ENUM，并返回查询结果给本地ENUM。

（5）主叫IMS网络S-CSCF/IBCF根据本地ENUM返回的结果，将业务请求路由到被叫归属的IMS网络。

通过该方案可以实现不同VoLTE之间的互通，并且可以实现VoLTE用户通过IMS网络的互通，非VoLTE用户不增加IMS网络的信令负荷，仍然通过传统电路域与其他运营商互通，这样保护了运营商的已有投资，与此同时还可以减少运营商对IMS网络的投资。

该方案要求互通的运营商之间，开放ENUM查询接口。对于个别运营商，出于安全等因素考虑，如果不愿意开放接口又想实现VoLTE互通，可以采用本方案。本方案可以通过共享ENUM上静态预配置，然后直接给查询请求的本地ENUM构造成功响应，响应结果的SIP统一资源标识符（SIP URI）组成包括：userpart部分为号码，domain部分为归属运营商的域名，如再添加user=phone参数，这样主叫IMS网络同样可以根据ENUM返回的SIP URI路由并实现VoLTE互通。无论被叫是否是VoLTE用户，都会通过IMS网络路由到归属运营商网络，这样就会增加运营商IMS网络的信令负荷，不利于保护运营商已有的电路域投资。

2.2.2 方案2

在VoLTE网络不断发展的后期，IMS用户不断增加，CS域用户不断减少，此时不光有VoLTE互通需求，还有IMS网络承接非IMS号码路由的需求。那么可以通过AS直接查询MNP的方案实现VoLTE与VoLTE的互通。图5是该方案的示意图，主要流程如下：

（1）主叫IMS网络S-CSCF网元在做完主叫业务后，主叫侧最后再次触发AS，如多媒体电话（MMTEL）AS。

（2）AS通过MAP接口直接查询MNP，根据MNP返回的号码携带RN信息，能够在被叫号码前面插入不同的前缀。

（3）S-CSCF/IBCF可以根据号码前缀将业务请求路由到不同的归属运营商。

通过该方案同样可以实现VoLTE互通，且实现了非IMS用户由IMS网络路由的功能。电路域网元无需再查询MNP，对于MNP没有增加负荷消耗。如果归属运营商不支持VoLTE网络，那么S-CSCF根据号码前缀路由到MGCF，然后由电路域路由到归属网络。

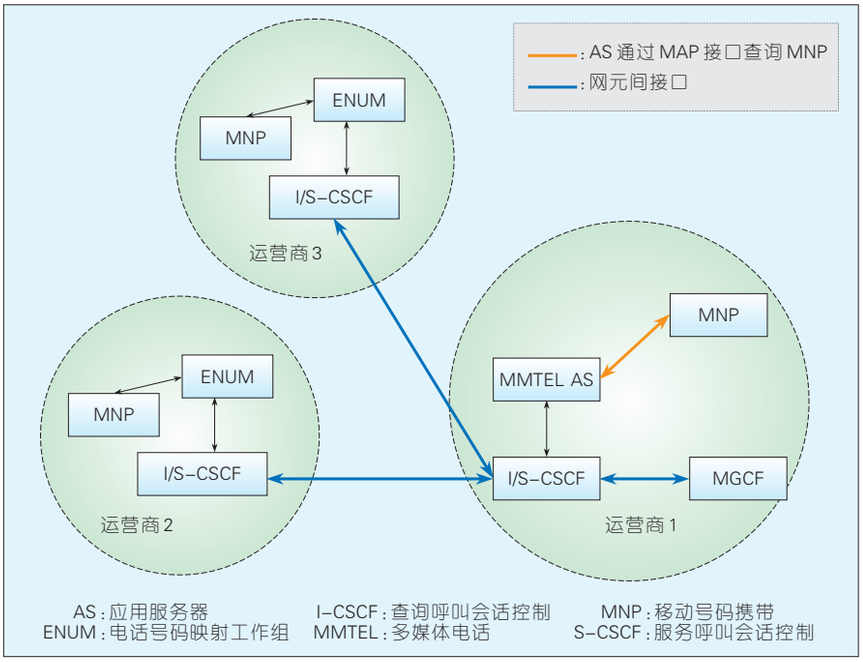
3 VoLTE互通部署建议

对于不支持号码携带业务的国家/地区，通过号码/号段信息能够知道被叫号码所归属的运营商信息，现有方案能够实现不同运营商之间的VoLTE互通。

对于支持号码携带业务的国家和地区，我们提出了几种有效解决方案，能够实现支持号码携带的不同运营商VoLTE网络之间互通。各种方案优、缺点和部署建议，见表1。

4 结束语

我们重点研究了支持号码携带业务的国家和地区中不同运营商VoLTE网络之间的互通要求，并提出



▲图5 MMTEL AS查询MNP方案

▼表1 VoLTE互通方案分析

适用场景	方案	优点	缺点	部署建议
两个VoLTE网络互通	方案1	实现互通，CSCF无需改动	ENUM需要定制逻辑，要求对方网络开放ENUM查询接口	适用两个VoLTE网络之间互通
	方案2	实现互通，ENUM无需改动	S-CSCF需要定制逻辑，要求对方网络开放ENUM查询接口	适用两个VoLTE网络之间互通
多个VoLTE网络互通	方案3	对IMS和CS网络改动较小，保护了运营商电路域已有投资	MNP和ENUM之间同步方案实时性不足，需要部署共享ENUM	适用于VoLTE网络建设初期
	方案4	仅需要MMTEL AS改动，对互通网络无特殊要求	IMS网络需要查询MNP，IMS网络需要承担非IMS用户的路由	适用于VoLTE网络后期，IMS用户较多场景

AS：应用服务器　ENUM：电话号码映射工作组　MNP：移动号码携带
CS：电路交换域　IMS：IP多媒体子系统　S-CSCF：服务呼叫会话控制
CSCF：呼叫会话控制　MMTEL：多媒体电话　VoLTE：基于LTE的语音业务

了4种有效解决方案，实现了不同运营商VoLTE网络之间互通。目前部分方案已经在香港地区进行了验证，方案可行，用户业务体验良好。这对于全球VoLTE网络互通有较强的借鉴意义。

参考文献
[1] 李中科, 梁斌, 廖芳芳. VoLTE和CS语音补充业务一致性方案研究[J]. 邮电设计技术, 2014(2): 6–10
[2] 朱斌, 文涛, 符刚. VoLTE部署关键问题研究[J]. 邮电设计技术, 2014(2): 1–5
[3] 刘扬, 张秀莹. 号码携带业务对IMS网络互通影响分析[J]. 邮电设计技术, 2012(5) : 29–32
[4] 王志松, 刘冬梅, 张林林. VoLTE中的业务域部署方案[J]. 移动通信, 2014(3): 57–62
[5] 缪永生. 通用引导架构在IMS网络中应用研究[J]. 中兴通讯技术, 2014, 20(4): 40–43.doi: 10.3969/j.issn.1009–6868.2014.04.010
[6] 许慕鸿. LTE语音目标解决方案_VoLTE技术[J]. 现代电信科技, 2013(11) : 33–45
[7] 3GPP TS 23.228 IP Multimedia Subsystem (IMS) [S]
[8] 3GPP TS 24.229 IP Multimedia Call Control based on SIP and SDP[S]
[9] 3GPP TS 23.221 Architectural Requirements Stage 1 [S]
[10] 3GPP TS 22.228 Service Requirements for the IP Multimedia Core Network Subsystem Stage 1 [S]

作者简介

缪永生，中兴软件有限责任公司系统工程师；研究方向为移动通信、IMS、VoLTE、虚拟仪器；已申请/授权15项专利。

倪明，中兴通讯股份有限公司IMS产品总工；研究方向为NGN、IMS、VoLTE；已发布10多篇论文、多项授权专利。

综合信息

2016年Wi-Fi设备出货量将新增30亿台

Wi-Fi联盟近日宣布WiFi设备出货量已经达到120亿台，到2016年年底有望冲破150亿台。

市场研究机构ABI Research研究主管菲尔·索利斯表示：随着2016年Wi-Fi设备出货量有望新增30亿台，运行在2.4 GHz和5 GHz频带的双频带设备出货量也将逐年增加，Wi-Fi增长势头毫无放缓的迹象。

Wi-Fi联盟还公布了2016年的技术路线图。同时，还将引进WiGig认证，支持多重视流影像、4K视频同步等多种情景模式。

此外，Wi-Fi位置计划将催生基于位置信息的新应用程序。在位置网络的覆盖范围内，支持Wi-Fi位置的设备将可对室内和室外的位置作出精确定位。

（转载自《中国信息产业网》）

DOI:10.3969/j.issn.1009-6868.2016.01.015
网络出版地址:http://www.cnki.net/kcms/detail/34.1228.TN.20151229.1509.002.html

辅助北斗技术的捕获空间计算和误差分析

Search Space Compute and Error Analysis of A-Beidou Acquisition Technology

中图分类号:TN929.5 文献标志码:A 文章编号:1009-6868 (2016) 01-0059-004

谢棋军/ XIE Qijun
陈新/ CHEN Xin
刘佩林/ LIU Peilin
(上海交通大学 电子工程系,上海 200240)
(Department of Electronic Engineering, Shanghai Jiaotong University, Shanghai 200240, China)

摘要: 结合辅助北斗定位技术的辅助信息类型,提出了北斗GEO和NGEO卫星在星历辅助条件下的卫星信号载波频率和码相位的估计方法,在此基础上重点研究了估计得到的载波频率和码相位的不确定度,并针对粗时间和精时间这两种典型的辅助定位技术进行了算例分析。研究结果表明:粗时和精时辅助能有效减少捕获搜索空间;引起载波频率搜索范围不确定度的因素是时钟频偏和接收机动态,而引起码相位搜索范围不确定度的因素是辅助时间精度和辅助位置误差。

关键词: 辅助定位技术;粗/精时间辅助;载波频率和码相位;捕获搜索范围

Abstract: With the help of assisted data of the A-BeiDou system, a method is proposed to estimate the shift of carrier frequency and pseudo-code phase of the Geostationary Earth Orbit (GEO) and Non-Geostationary Earth Orbit (NGEO) satellites. We analyze the uncertainty of the estimated value of carrier frequency and pseudo-code phase. Finally, two typical assistant positioning technologies—coarse time and fine time assisted positioning—are analyzed. Simulation shows that assisted data reduces the acquisition searching space; the clock bias and receiver dynamic are two important factors that cause the error of carrier frequency; and assist time precision and assist position error are two key factors that led to the error of code phase.

Key words: tassisted positioning technology; coarse time and fine time assisted; carrier frequency and pseudo-code phase; search range capture

随着导航技术的发展以及实际应用对于弱信号快速定位的需求,辅助全球导航卫星系统(A-GNSS)[1-3]技术已经成为导航领域的研究热点。目前在其他国家,卫星定位系统(GPS)、GALILEO、GLONASS等导航系统均已进入第三代合作伙伴计划(3GPP)组织并形成相应的A-GNSS行业标准,而中国辅助北斗定位技术(A-Beidou)的标准化工作处于起步阶段,亟需针对其特点对各项指标进行不断研究,推进其加入3GPP标准。

收稿日期:2015-11-02
网络出版时间:2015-12-29
基金项目:国家自然科学基金(61304225)

1 A-Beidou捕获搜索空间的估计

卫星信号捕获是一个频率和码延迟的二维搜索过程,卫星和接收机在两者连线方向上的相对运动所引起的多普勒效应决定了接收机在载波频率一维空间搜索的范围,接收机接收到卫星信号的时刻决定了接收信号的伪码相位值,时刻的精度决定了接收机在伪码相位空间内搜索的范围。

1.1 载波频率的估计

利用辅助星历/历书、时间可以计算出卫星的位置和速度[4],如果接收机的位置和速度也已知,则可以进一步计算出卫星和接收机连线方向的多普勒频偏,为捕获提供初始信息。北斗卫星包括地球同步轨道卫星(GEO)和非地球同步轨道卫星(NGEO),卫星的位置和速度需要分开计算。接收机接收到的北斗卫星信号的多普勒频移计算公式如下:

$$f_d = f\frac{\|\dot{v}^{(s)} \cdot e^{(s)}\|}{c} \quad (1)$$

$$\dot{v}^{(s)} = \dot{v} - \dot{v}_{rec} \quad (2)$$

$$e^{(s)} = \frac{(X_k^{(s)}\ Y_k^{(s)}\ Z_k^{(s)}) - (X_{rec}\ Y_{rec}\ Z_{rec})}{\|(X_k^{(s)}\ Y_k^{(s)}\ Z_k^{(s)}) - (X_{rec}\ Y_{rec}\ Z_{rec})\|} \quad (3)$$

其中,北斗B1I载波频率 f = 1561.098 MHz, $\dot{v}^{(s)}$ 为第 s 颗北斗卫星速度矢量与接收机速度矢量差, c 为光速, $e^{(s)}$ 为信号在发射时刻为 t_k 时接收机至卫星的单位观测矢量, $(X_{rec}, Y_{rec}, Z_{rec})$ 为 t_k 时刻接收机坐标,

$(X_k^{(s)}\ Y_k^{(s)}\ Z_k^{(s)})$ 为 t_k 时刻第 s 颗 NGEO 卫星在 CGCS 2000 坐标系下的坐标。

1.2 码相位的估计

北斗卫星信号的码相位由卫星传输时间所决定，第 k 颗卫星传输时间为：

$$TOT^k = SOW^k + \Delta T^k \tag{4}$$

$$\Delta T^k = N_{bit}^k + N_{ms}^k + N_{codephase}^k \tag{5}$$

其中，N_{bit}^k 代表第 k 颗卫星在找到 SOW^k 之后导航比特的个数，N_{ms}^k 代表第 k 颗北斗卫星传输完最后一个导航电文比特之后的整毫秒数，也称为扩频码周期数，$N_{codephase}^k$ 为第 k 颗北斗卫星的码相位。N_{bit}^k、N_{ms}^k 和 $N_{codephase}^k$ 以时间为单位[5]。

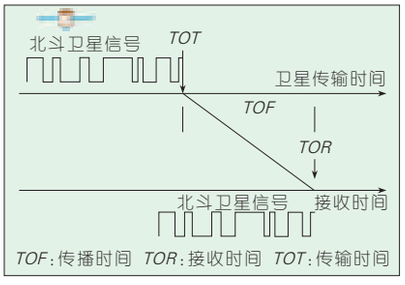
辅助北斗定位技术以接收时间作为辅助的时间信息，根据辅助的接收时间和接收机以及通过辅助星历/历书计算得到的卫星坐标，可以反推出北斗卫星传输时间，并可以计算得到码相位。如图 1 所示，北斗卫星的传输时间记为 TOT，接收时间为 TOR，卫星信号传播时间记为 TOF，则码相位 $N_{codephase}$ 的计算方法为：

$$N_{codephase} = \mathrm{mod}(TOT, 1\ ms) * V_{code} \tag{6}$$

$$TOT = TOR - TOF \tag{7}$$

$$TOF = \sqrt{(X_s - X_{rec})^2 + (Y_s - Y_{rec})^2 + (Z_s - Z_{rec})^2}/c \tag{8}$$

其中 (X_s, Y_s, Z_s) 为由星历/历书估算的卫星位置，$(X_{rec}, Y_{rec}, Z_{rec})$ 为接收机参考位置，V_{code} 为北斗码片速率，为 2.046 Mbit/s。需要人们注意的是：由于北斗信号一个扩频码周期为 1 ms，因此，当时间辅助信号精度高于 1 ms 才能对码相位进行估计。



▲图 1 北斗卫星传输时间和接收时间

2 A-Beidou 捕获搜索空间模糊度分析

2.1 频域搜索空间模糊度分析

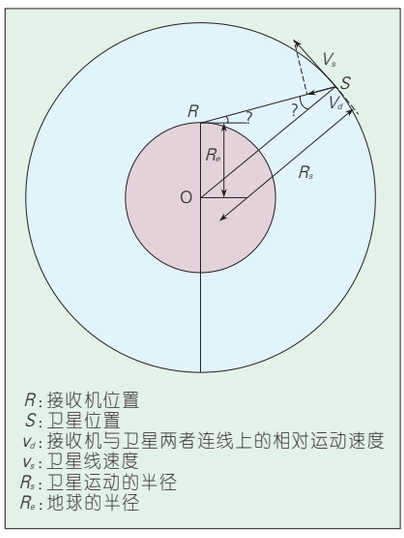
（1）辅助时间精度对载波频率的影响

假设北斗导航系统（BDS）卫星轨道是一个圆，接收机在卫星轨道平面和地球表面相交的大圆上，将轨道平面放平，得到图 2。在图 2 中，S 为卫星位置，R 为接收机位置，O 为地球球心，v_s 为卫星线速度，v_d 为接收机与卫星两者连线上的相对运动速度。

如图 2，根据几何关系，有：

$$v_d = \frac{v_s R_e \cos\theta}{R_s} \tag{9}$$

其中，v_s 表示在地心地固坐标系下的卫星运动速度，GEO 卫星速度认为是 0，IGSO 卫星速度为 1 ~ 3 km/s，MEO 卫星速度为 2 ~ 4 km/s；R_e 表示地球的半径，这里假设取平均半径 6 371 km；θ 表示接收机和卫星之间的仰角，在 0° ~ 180° 之间；R_s 为卫星到地心的距离，GEO 和 IGSO 卫星轨道高度为 35 786 km，因此它们的 R_s 为 42 157 km，MEO 卫星轨道高度则可以达到 21 528 km，其 R_s 为 27 899 km。



▲图 2 卫星和接收机运动分析

当 θ =90° 的时候，v_d 最小为 0；当 θ = 0° 或 180° 时，可以获得 v_d 的最大值。对于 IGSO 卫星，多普勒变化率的计算公式为：

$$\frac{df_d}{dt} = \frac{d(\frac{v_d}{c} f_{B1})}{dt} = \frac{dv_d}{dt} \cdot \frac{f_{B1}}{c} = \frac{dv_d}{d\theta} \cdot \frac{d\theta}{dt} \cdot \frac{f_{B1}}{c} \tag{10}$$

因此，当 θ =0° 或 180° 时，$\frac{df_d}{dt}$ 取 0 Hz/s；θ =90° 时，$\frac{df_d}{dt}$ 取最大值（这里指绝对值），带入数值，计算便可得到北斗卫星的最大多普勒变化率为 0.896 Hz/s。因此，辅助时间造成的多普勒不确定度 Δf 为：

$$\Delta f = 0.896 \cdot \Delta t \tag{11}$$

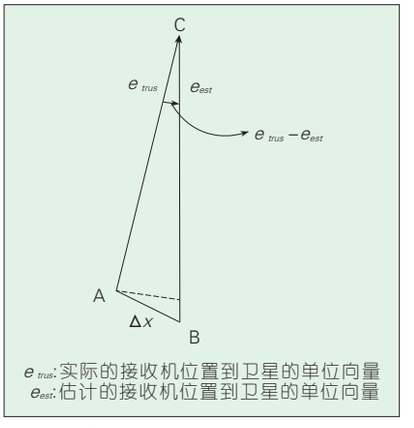
其中，Δt 为辅助时间的精度。

（2）接收机位置误差对载波频率的影响

辅助位置会对估算的多普勒值产生影响。如图 3 所示，A 为接收机的准确坐标，B 为接收机的估计坐标，C 为卫星的坐标，卫星多普勒等于速度向量点乘接收机和卫星连线方向上的单位向量，卫星的多普勒误差为：

$$\begin{aligned}\Delta f &= B1I/c \cdot (-v \cdot e_{est} + v \cdot e_{true}) \\ &\approx B1I/c \cdot v \cdot (e_{true} - e_{est})\end{aligned} \tag{12}$$

该方程最后一个等式取约等于是由于 e_{true}、e_{est} 与卫星速度之间的夹角稍有不同，这两个夹角由于卫星



▲图 3 辅助位置误差

到接收机的距离远大于接收机的误差距离，因此近似相等。当B和D重合并且卫星速度和位置误差向量方向垂直时，多普勒误差达到最大值：

$$\Delta f_{\max} = B1I/c \cdot v \cdot |\Delta x|/CD \quad (13)$$

根据北斗空间接口文件，对于GEO和IGSO卫星，CD取36 786 km；对于MGEO卫星，CD取21 528 km。

（3）接收机速度和钟差对载波频率的影响

假设接收机向着卫星以km/h的速度运动，速度引起多普勒偏移量为：

$$\Delta f = B1I \cdot v/c \quad (14)$$

而对于未知的振荡器频偏，可知每N ppm的时钟频偏产生的未知载波频偏：

$$\Delta f = N \cdot B1I \cdot 10^{-6} \quad (15)$$

在式中（14）、（15）中，$B1I$为北斗$B1I$信号的标称载波频率，为1 561.098 MHz，c为光速。

2.2 码相位搜索空间模糊度分析

（1）辅助时间精度对码相位一些影响

由于北斗的一个扩频码的周期是1 ms，因此，只有当时间精度高于1 ms才可以缩小码相位的搜索范围。

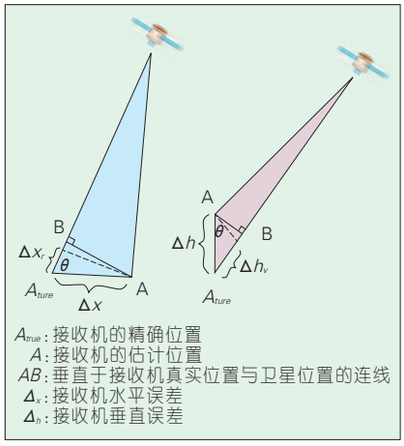
时间误差Δt造成的码相位误差$\Delta code$为：

$$\Delta code = \Delta t \cdot V_{code} \quad (16)$$

其中V_{code}为北斗扩频码速率，为2.046 Mcps。

（2）接收机位置误差对码相位的影响

在分析接收机位置误差对码相位的影响时，将接收机位置误差分解为垂直误差和水平误差，如图4所示，A_{true}为接收机的精确位置，A为接收机的估计位置，AB垂直于接收机真实位置与卫星位置的连线，Δx为接收机水平误差，Δh为接收机垂直误差，接收机水平和垂直误差造成的距离误差为：



▲图4 辅助位置水平和垂直误差

$$\Delta x_r = \cos(\theta) \cdot \Delta x \leq A_{true}B \quad (17)$$

$$\Delta h_v = \sin(\theta) \cdot \Delta h \leq A_{true}B \quad (18)$$

因此，接收机误差造成的总的距离误差ΔL为：

$$\Delta L = \Delta x_r + \Delta h_v \leq \cos(\theta) \cdot \Delta x + \sin(\theta) \cdot \Delta h \quad (19)$$

造成的码相位误差$\Delta code$为：

$$\Delta code = \Delta L \cdot V_{code}/c \quad (20)$$

其中V_{code}为北斗扩频码速率，为2.046 Mcps，c为光速。

（3）大气层延迟对码相位的影响

大气层延迟主要包括电离层延迟和对流层延迟两部分，其影响量级从几米到几十米，是全球卫星导航系统（GNSS）导航定位的主要误差源，恶劣情况下大气层延迟会对码相位产生多个码片的误差。

（a）电离层延迟估计

北斗电离层模型的建立基于地理坐标系，利用刺穿点的地理经度与太阳地理经度的差值和刺穿点的地理纬度作为变量构造电离层模型。北斗提供了Klobuchar8参数电离层改正模型来估计电离层延迟，得到电离层延迟$I_z(t)$。

（b）对流层延迟估计

现有的估算修正对流层延迟的模型有很多种，主要有Hopfield模型、Saastamoinen模型等，我们采用精度较好的Saastamoinen模型估算对流层延迟，得到电离层延迟$I_z(t)$。

（c）大气层延迟所造成的码相位误差

根据(a)、(b)的分析，大气层延迟对码相位造成的误差为：

$$\Delta code = [\Delta S + I_z(t)] \cdot V_{code}/c \quad (21)$$

其中V_{code}为北斗扩频码速率，为2.046 Mcps；c为光速；ΔS和$I_z(t)$为估计的对流层和电离层造成的距离误差，当误差为100 m时，将导致将近一个码片的误差。

3 A-Beidou捕获搜索空间计算和分析实例

本节将给出粗时和精时两种算例，分别估算北斗卫星的捕获空间，并对估算模糊度进行分析。

选择NGEO卫星6，采用的辅助数据包括以下几个参数。

（1）时间：2014年7月1日上午8点整，时间精度粗时辅助为2 s，精时辅助为10 us。

（2）星历：2014年7月1日上午8点整的1号卫星和6号卫星的16参数星历。

（3）接收机位置：地理坐标为北纬121.4°、东经31.2°、高度为6 m，接收机位置水平误差为5 km，接收机速度80 km/h。

(4）接收机晶振频偏：100 ppb。

在精时辅助下6号卫星的速度为1 315.8 m/s，进一步求得载波频率估计值为243.46 Hz，根据公式（6）—（8）求得码相位估计值为1 108.84码片。根据频率搜索空间模糊度分析公式（11）、（13）—（15）可以分别计算得到粗时辅助和精时辅助频率误差的结果，见表1、表2。根据公式（16）、（20）、（21）得到精时辅助的码相位误差结果，见表3。

3.1 粗时间辅助搜索空间分析

根据计算结果我们可以知道：6

▼表1 北斗6号卫星的频率搜索各分量误差

辅助参数	造成的频率搜索误差/Hz	占搜索空间百分比/%
辅助时间 ± 2 s	± 1.792	0.65
辅助位置 5 km	± 1.589	0.58
运动速度 80 km/h	± 116	42.11
晶振钟偏 100 ppb	± 156.11	56.67
频率搜索的总范围：± 275.491 Hz，频率估计值：243.36 Hz		

▼表2 北斗6号卫星的频率搜索各分量误差

辅助参数	造成的频率搜索误差/Hz	占搜索空间百分比/%
辅助时间 ± 10 us	± 0.00001	0.00
辅助位置 5 km	± 1.590	0.58
运动速度 80 km/h	± 116	42.46
晶振钟偏 100 ppb	± 156.11	57.04
注：频率搜索的总范围为 ± 273.7 Hz，频率估计值为 243.46 Hz		

▼表3 北斗6号卫星的码相位搜索各分量误差

辅助参数	造成的码相位误差/码片	占搜索空间百分比/%
辅助时间 ± 10 us	± 20.46	37.36
辅助位置 5 km	± 33.52	61.21
电离层延迟	± 0.17	0.31
对流层延迟	± 0.61	1.11
注：码相位搜索的总范围为 ± 54.76 码片，码相位估计值为 1 108.84 码片		

号卫星总的辅助频率搜索范围为 −275.491 ~ +275.491 Hz，同样的算例，不采用相关辅助信息，搜索范围将达到 −5 ~ +5 kHz，大概是采用辅助信息时搜索范围的 18.18 倍。晶振的偏差引起的频偏占整个捕获频率搜索总空间的很大一个比例。

3.2 精时辅助搜索空间分析

根据计算可知：6号卫星总辅助频率搜索范围为 −273.70 ~ +273.70 Hz，同样的算例，不采用精时辅助，搜索空间将小于采用粗时辅助下的搜索空间，也远远小于不采用辅助信息时 −5 ~ +5 kHz 的搜索空间，但是优化的搜索空间较粗时间辅助也就是减少了个位数的搜索 Hz 数。粗时间辅助对码相位搜索空间是没有变换的，都是 2 046 个码片。从表3可以看出精时间辅助可以大大缩小码相位的搜索空间。在本算例中6号卫星码相位搜索空间由不采用辅助信息的 2 046 个码片减少至 −54.76 ~ 54.76 码片，大概会减少了将近 19 倍的搜索空间。

4 结束语

结合辅助北斗定位技术的辅助信号类型，我们给出了北斗 GEO 和 NGEO 卫星在星历、粗/精时间和接收机位置辅助条件下的卫星信号载波频率和码相位的计算方法，并在此基础上重点研究了计算得到的载波频率和码相位的不确定度，最后针对粗时间和精时间这两种典型的辅助定位技术进行了算例分析。结果表明，在粗时间和精时间辅助前提下，北斗 GEO 卫星和 NGEO 卫星的载波频率搜索范围比没有辅助信息时的搜索范围减小了近 20 倍，接收机时钟的频偏和接收机的动态是引起捕获过程中频率搜索不确定度的两个主要因素，辅助位置精度对 NGEO 载波搜索空间较 GEO 载波搜索空间影响大。在精时间辅助前提下，捕获搜索的码相位较其他方式减小了近 20 倍。并认为引起码相位不确定度的两个重要因素是辅助时间和辅助接收机位置的精度。

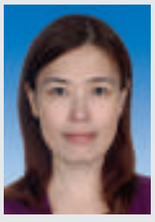
参考文献

[1] 汶晓勇, 肖越. GPS 和 A-GPS 技术研讨[J]. 通信技术, 2011, 44(8): 76–78
[2] 丁翔宇. GPS 卫星导航定位技术的新进展[J]. 全球定位系统, 2008, 33(3): 46–49
[3] 严昆仑, 章红平, 张提升, 等. AGPS 系统原型设计与性能评估[C]//第四届中国卫星导航学术年会论文集–S5 卫星导航增强与完好性监测, 2013
[4] 李显, 吴美平, 张开东, 等. 导航卫星速度和加速度的计算方法及精度分析[J]. 测绘学报, 2012, 41(6): 816–824
[5] VAN D F, ABRAHAM C. Coarse-Time AGPS; Computing TOW From Pseudorange Measurements, and the Effect on HDOP[C]// in Proceedings of the 20th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2007), 2001: 357–367

作者简介

谢棋军，上海交通大学电子科学与技术硕士研究生；主要研究方向为北斗导航和辅助型定位的 SOC 技术等；发表论文3篇，获得专利1项。

陈新，上海交通大学北斗导航与位置服务上海市重点实验室助理研究员；主要研究方向为导航高灵敏度高精度相关算法；参与负责中国航天科技集团卫星应用技术研究创新基金、国家自然科学基金青年项目和中国第二代卫星导航系统重大专项等专项项目；近年来发表论文7篇，获得专利2项。



刘佩林，上海交通大学教授，上海交通大学空间信息技术研究中心副主任；主要研究方向为导航、通信及低功耗高性能 SoC 架构、导航与空间 SoC 技术发展方向等；已发表篇论文50多篇，获得专利30余项。

《中兴通讯技术》杂志(双月刊)投稿须知

一、杂志定位

《中兴通讯技术》杂志为通信技术类学术期刊,通过介绍、探讨通信热点技术,展现通信技术最新发展动态,并促进产学研合作,发掘和培养优秀人才,为振兴民族通信产业做贡献。

二、稿件基本要求

1. 投稿约定

(1) 作者需登录《中兴通讯技术》投稿平台:www.zte.com.cn/paper,并上传稿件。第一次投稿需完成新用户注册。

(2) 编辑部将按照审稿流程聘请专家审稿,并根据审稿意见,公平、公正地录用稿件。审稿过程需要1个月左右。

2. 内容和格式要求

(1) 稿件须具有创新性、学术性、规范性和可读性。

(2) 稿件需采用WORD文档格式。

(3) 稿件篇幅一般不超过6 000字(包括文、图),内容包括:题名、作者姓名、作者单位、中文摘要、关键词(4 ~ 8个)、英文摘要、正文、参考文献、作者简介。

(4) 中文题名一般不超过20个汉字,中、英文题名含义应一致。

(5) 摘要尽量写成报道性摘要,包括研究的目的、方法、结果与结论,以150 ~ 200字为宜。摘要应具有独立性和自明性,采用第三人称。中英文摘要内容应一致。

(6) 文稿中的量和单位应符合国家和国际标准。外文字母的正斜体、大小写等须写清楚,上下角的字母、数据和符号的位置皆应明显区别。

(7) 图、表力求少而精(以8幅为上限),应随文出现,切忌与文字重复。图、表应保持自明性,图中缩略词和英文均要在图中加中文解释。表应采用三线表,表中缩略词和英文均要在表内加中文解释。

(8) 参考文献以20条左右为宜,未公开发表的资料不宜列入。所有文献必须在正文中引用,文献序号按其在文中出现的先后次序编排。主要种类参考文献的书写格式为:

- 期刊[序号]作者. 题名[J]. 刊名,出版年,卷号(期号):起止页码
- 书籍[序号]作者. 书名[M]. 出版地:出版者,出版年:起止页码
- 论文集中析出文献[序号]作者. 题名[C]//论文集编者. 论文集名(会议名). 出版地:出版者,出版年(开会年):起止页码
- 学位论文[序号]作者. 题名[D].地点:学位授予单位,授予年
- 专利[序号]专利所有者. 专利题名. 国别:专利号[P]. 公布日期
- 国际、国家标准[序号]标准编号,标准名称[S]

(9) 作者原则上不超过3人,超过3人时,可以感谢形式在文中提及。作者简介包括:姓名、工作单位、职务或职称、学历、毕业于何校、现从事的工作、专业特长、科研成果、已发表的论文数量等。

(10) 提供正面、免冠、彩色标准数码照片一张,最好采用JPG格式(文件大小超过100 kB)。

(11) 尽可能标注出研究课题的资助基金或资助项目名称。

(12) 作者姓名中含有多音字时,应标注作者姓名的汉语拼音。

(13) 提供联系方式,如:通信地址、电话(含手机)、Email等。

3. 其他事项

(1) 请勿一稿多投。凡在2个月(自来稿之日算起)以内未接到录用通知者,可致电编辑部询问。

(2) 为了促进信息传播,加强学术交流,在论文发表后,本刊享有文章的版权(包括英文版、电子版、网络版和优先数字出版)。作者获得的稿费包括版权酬金。如对此持有不同意见,请在投稿时说明。

编辑部地址:安徽省合肥市金寨路329号国轩凯旋大厦1201室,邮政编码:230061

联系电话:0551-65533356,联系邮箱:magazine@zte.com.cn

本刊只接受在线投稿,欢迎访问本刊投稿平台:www.zte.com.cn/paper

中兴通讯技术
ZHONGXING TONGXUN JISHU

双月刊 1995年创刊 总第126期
2016年2月 第22卷第1期

主管：安徽省科学技术厅
主办：安徽省科学技术情报研究所
中兴通讯股份有限公司
编辑：《中兴通讯技术》编辑部

《中兴通讯技术》编辑部
地址：合肥市金寨路329号凯旋大厦12楼
邮编：230061
网址：www.zte.com.cn/magazine
投稿平台：www.zte.com.cn/paper
电子信箱：magazine@zte.com.cn
电话：(0551)65533356
传真：(0551)65850139

出版、发行：中兴通讯技术杂志社
发行范围：全球发行
印刷：合肥添彩包装有限公司
出版日期：2016年2月10日
刊号：ISSN 1009-6868
CN 34-1228/TN
广告经营许可证：皖合工商广字0058
定价：每册20.00元，全年120.00元