大数据安全

丁 锋 单世民 兰 兰 郭剑锋◎著

大数据概念在研究和应用领域都已经大行其道，大数据安全成了无法回避的现实问题。本书对大数据安全的现状、技术、策略进行了梳理，从安全视角讨论和分析大数据，努力使读者对大数据安全形成一个完整的轮廓。

中国言实出版社

图书在版编目（CIP）数据

大数据安全 / 丁锋等著 . -- 北京：中国言实出版社，2016.5
ISBN 978-7-5171-1888-6

Ⅰ. ①大… Ⅱ. ①丁… Ⅲ. ①数据处理—安全技术
Ⅳ. ① TP274

中国版本图书馆 CIP 数据核字（2016）第 103261 号

责任编辑：邓见柏
文字编辑：李　琳
封面设计：王立霞

出版发行　中国言实出版社
地　址：北京市朝阳区北苑路 180 号加利大厦 5 号楼 105 室
邮　编：100101
编辑部：北京市海淀区北太平庄路甲 1 号
邮　编：100088
电　话：64924853（总编室）　64924716（发行部）
网　址：www.zgyscbs.cn
E-mail：zgyscbs@263.net
经　　销　新华书店
印　　刷　北京京华虎彩印刷有限公司
版　　次　2016 年 5 月第 1 版　　2016 年 5 月第 1 次印刷
规　　格　710 毫米 ×1000 毫米　1/16　6 印张
字　　数　86 千字
定　　价　16.00 元　　ISBN 978-7-5171-1888-6

CONTENTS

目录

第 1 章　大数据安全现状

大数据技术的发展为数据价值的发掘提供了舞台，也引发了新一轮的数据安全与隐私保护问题。

大数据技术的应用在生活中随处可见。例如当用户通过微信扫描二维码并转发信息时，大数据分析工具会捕捉到用户的消费习惯及个人喜好，同时对用户需求进行分析和预测，通过分析结果为用户提供更多服务，公众知情或不知情的情况下提供了数据，于是安全问题也由此产生。

现在是大数据发展的重要时期。信息安全是大数据发展过程中无法回避的巨大挑战。很多消费者在不知情的情况下被相关公司搜集、窃取到了个人信息。更令人担忧的是，中国尚未出台个人信息保护法，只有部分法律法规中零散提及个人信息安全。因为没有上位法[①]，很多与大数据相关的活动的合法性便无从说起。目前只能希望企业在运用大数据技术获得利润的同时重视信息安全问题，能够做好相应的防范与保护措施，保护消费者的隐私。

（1）大数据应用日渐广泛，带来诸多安全问题

当用户下载手机应用的时候，往往会弹出是否允许该应用共享用户的通讯及位置信息的询问框，部分用户会选择允许该要求，这就意味着该用户将面临信息泄露的风险，原因在于大数据时代，用户是无法阻止外部数据商获取个人信息的。大量用户的信息及各项数据会被社交、购物等网站对外开放，而这些信息又会被同类网站或市场分析机构收集，通过分析用户的个人信息、手机定位等多种数据信息的拼凑形成的数据集合，就能够分析用户的信息体系，这就使用户的隐私处在危险当中，令人担忧。

---

① 上位法：就法的效力位阶而言，法可分为三类，即上位法、下位法和同位法。就法律效力大小而言，效力大的为上位法，它之下生效的为下位法。比如说宪法和其他法律部门的关系，宪法就是上位法，因为其他法律都是依据宪法制定的，其他的法律如刑法、民法就是下位法。

如果用户拒绝共享信息，那将导致其无法享受到部分便捷的服务，这种结果说明互联网的发展正逐步变得更加依赖公民的个人信息。而对公民信息需求的增长及获取渠道的拓宽导致了信息安全、隐私及便利性之间产生巨大的冲突。一方面，消费者从海量数据中获益，这种益处包括低廉的价格、更符合消费者消费习惯和喜好的商品、生活品质的提高等。另一方面，厂商大量获取消费者的购买偏好及健康数据给用户的隐私安全带来威胁。

（2）进入大数据时代，企业面临多重安全风险

对于企业而言，在大数据迅猛发展的环境中不仅要学会如何利用数据挖掘价值，促进利益增长，还需要统筹安全部署，制定预案，防御数据泄露以及网络攻击。

高德纳咨询公司[①]曾提出："大数据安全是一场必要的斗争。"企业可以利用数据分析及挖掘获取利润，黑客也可以通过同样的方式向企业或个人发起攻击，而大数据技术无疑为黑客进行更精确的攻击提供了帮助。比如黑客可以先在社交网络、邮件、电子商务网站中获取攻击对象的电话、家庭住址等个人信息，然后当获取攻击对象的 VPN 账号后，黑客就可以得到攻击对象的工作信息，进一步侵入企业网络。对大数据分析有很高要求的企业面临的挑战更多。比如金融及天气预报的分析预测、复杂网络计算和广域网感知等，对于这些机构来说，传统安全防护很难达到效果，因为大数据安全和大数据业务是相对应的，任何一个会导致目标信息的提取和检索方向出现错误的攻击都会对企业大数据分析产生误导，检测方向也会出现偏差。而这些攻击需要企业集合大量数据并对其进行关联分析才能明白其攻击目的。因此大数据安全要求企业先要对自身的业务需求进行分析，针对可能威胁到大数据业务的因素提出预案。

（3）大数据时代，国家安全需直面信息战与网络恐怖主义

从国家安全层面看，信息时代与工业时代的不同之处就在于，仅凭军事防御已经不能够使国家各种重要设施及信息机构免受打击破坏，安全环

① 高德纳咨询公司：*Gartner* Group，成立于 1979 年，它是第一家信息技术研究和分析公司。它为有需要的技术用户提供专门的服务。

境发生了质的改变。如今网络脆弱，攻击者增加，攻击技术不断变化，国家在水、电、石油、商业、交通、军事等领域对网络的依赖性不断增加，使得国家安全面临巨大挑战。

对于网络恐怖主义来说，在大数据时代，庞大且全面的用户数据无疑为其提供了新的资源支持，入侵民众生活的各个方面变得轻而易举。美国政府为了更好地应对网络恐怖主义的袭击，决定建立个人信息库，通过商业手段收集涵盖美国民众金融、消费、旅游等各行各业的数据，为国家安全服务。这种措施也并非是全新的方式，早在“9·11”[①]事件之前就有趋势，并随着大数据的应用逐步增强。在新的数据环境下，对个人信息识别趋向更深和更广，对数据分布形势的分析能力也在不断提高。

对于信息的保护，公民自身和公共权力机构都要行动起来。

公民的信息安全保护意识是信息安全保护的第一道防线。单纯依赖政府部门与司法机关等权力机构去保护个人信息安全的效果有限，唤醒公民信息安全意识是预防信息泄露的重要途径。

当然，执法保护也发挥着十分重要的作用，但属于第二道防线。而在我国，执法机关对于互联网信息的保护处于初级阶段，存在着办理经验少、处理不及时、案件分类不明确等问题。据不完全统计，自 2003 年起，在我国有关互联网案件的判决数量不超过 150 件，由此可见执法及司法机关介入之少，保护互联网信息需要权力机关更加努力。

针对执法及司法机关存在的问题，加强立法被认为是治本之策。在互联网竞争中，对于其秩序规则的制定，既需要专门互联网方面的立法，同时也需要传统立法与之相扶持。如果只存在传统立法，对于专门性事件就会缺乏同等的约束力。例如我国的《反不正当竞争法》《反垄断法》已经制定了很长时间，这些法律虽然对竞争和垄断关系起到了一定的管束作用，但是由于没有制定专门的互联网规则，导制其对互联网的约束作用微乎其微。而如果没有传统立法做基础，互联网立法也就失去了存在的土壤，很多根本性问题便无法规范。

---

① “9·11 事件”：“9·11 事件”（又称“9·11”“9·11 恐怖袭击事件”），是 2001 年 9 月 11 日发生在美国本土的一起系列恐怖袭击事件。

因此，只有将传统法律与互联网专门规则相结合，才能同时顾全普遍性与专门性，才能真正提供有效的秩序规范。除此之外，法律规范与指导性规范共同施行也不失为一种维护互联网秩序的有效方法。

1. 大数据安全事件

在《2014 年度数据泄露调查报告》中，威瑞森公司[①]列举了 63737 起网络安全事件和已经确认出现的 1367 起数据泄漏事件。从报告中可以了解到，高达 25% 的信息泄露源于数据库，这其中深层的技术原因值得探寻。除此之外还有很多未确认、未公开或尚在调查的信息泄露事件没有公布。

2014 年国内外发生了几起重大的信息泄漏事件：

（1）春运第一天 12306 网站泄露用户信息

2014 年铁路春运售票第一天，12306 网站便多次出现程序漏洞，导致大量用户信息泄露。在早上经历了 1 小时的宕机后，12306 铁路客户服务中心网站又出现用户账号串号的问题，大量用户身份证等信息遭泄露。当天 15 时左右，用户在登录后能够查看到其他用户的个人信息，包括姓名、身份证号码、手机号码等。半个小时后，新版 12306 网站出现危害等级很高的资料泄露漏洞。

（2）前支付宝员工贩卖 20G 用户资料

前支付宝员工通过后台程序下载了超过 20G 的数据信息，并将其贩卖给部分商业公司及数据公司，其中价值高的单条数据记录价格甚至达到数十元。此次事件引起了大众对信息安全问题的关注，也牵出了一条互联网信息贩卖的黑色交易链。

在信息贩卖产业中，有很多种信息交易类型，有的通过公司运作方式，将从互联网中购买到的户籍、房产等个人信息贩卖给提供婚恋、定位等服务项目的公司获取利润；或者直接服务自身，通过窃取用户网上购买记录推销相应产品；等等。为了防止数据流入黑市，电商企业需要对自身

① 威瑞森公司：Verizon，美国最大的本地电话公司、最大的无线通信公司，全世界最大的印刷黄页和在线黄页信息的提供商。

的数据库系统进行严密监控。

（3）2000 万开房信息泄露案开庭

一个名为“查开房”的网址被实名认证账户“@ 股社区”发布在新浪微博中。在这个网址中，只需输入身份证号或姓名便可查询到包括身份证号码、家庭住址、出生日期、电话号码、公司、登记日期等个人隐私。自“开房数据泄露”事件发生后，用户饱受短信推销广告的骚扰，隐私权受到严重侵害。

（4）软件商变身“黑客”删除车管所内万余条违章记录

案件起源于某软件供应商与“黄牛”勾结并收买专业人员，通过车管所系统中的运维技术功能提供的便利，将事先编辑好的删除软件植入到车管所的软件系统中，同时更改了公安系统的网络配置，使其删除操作完美地避开了公安内网报警系统及现场监管，非法删除了共计一万四千余条的交通违章记录。经公安机关查明，这起案件涉案金额共计 1800 余万元。

（5）国内 130 万考研用户信息被黑产[①]利用

截止到 2014 年 11 月，国内共计有 130 万考研生的报名信息遭到泄露并被贩卖，泄露内容包括用户姓名、电话号码、家庭住址、身份证号、所在学校及专业等私密信息。同时有报道称漏洞平台遭到泄露的信息正在被黑产利用，危害巨大。

（6）韩国上亿条用户信用卡个人信息遭泄，引发信息贩卖潮

此次事件可称为韩国史上规模最大的信用卡信息泄露事件。信用评级公司内部某职员在受信用卡公司委托，开发电脑程序过程中，利用职务便利非法窃取了高达 1.04 亿的用户信息，包括姓名、电话号码、居住地址、身份证号码甚至贷款交易内容及信用卡免税书等各项敏感信息。此次大规模用户信息泄露直接导致了 KB 国民卡、乐天卡及 NH 农协卡公司社长引咎辞职。信用卡信息泄露的严重性不仅在于私密信息遭到窃取，更在于通过对信用信息的分析能够了解到顾客的消费能力及习惯，从而为不法分子

① 黑产：所谓黑色产业（简称黑产），就是利用病毒木马来获得利益的一个行业。由于干的是龌龊的勾当，故称黑色产业。专指电脑网络，如果引申，贩毒、高利贷、私彩、短信诱骗、语聊之类也可以算黑色产业。故黑色产业就是利用非法的手段获取利益的龌龊勾当。

提供金融欺诈或强制消费的机会。

（7）土耳其黑客侵入本国电力系统，删除巨额债务记录

RedHack（土耳其黑客组织）上传了其在电力管理系统中更改及删除债务数据的视频。在此次入侵中，RedHack 撤销了索玛（Soma）地区需要交付给电力公司的 150 万土耳其币（约合 65 万美元）的巨额债务账单。同时他们还公布了该系统的登录用户名及密码。

（8）中国互联网 DNS 服务器发生长时间大面积瘫痪

DNS（域名解析系统）是能够将网站域名翻译成一串数字的互联网基础设施。2014 年 1 月 21 日，互联网 DNS 出现事故，该故障造成了近三分之二的服务器瘫痪，持续时间长达数小时，使 85% 的互联网用户受到波及无法正常使用网络。

（9）OpenSSL[①] 发现心脏出血漏洞

多起 OpenSSL 的“Heartbleed”心脏出血漏洞攻击事件相继发生并公布，例如访问量很高的淘宝、支付宝、微信公众号、网银、陌陌等基本上都出现了心脏出血漏洞，而这些漏洞能够帮助黑客十分容易地获得用户的 Cookie[②] 甚至明文账号与密码等私密信息。

（10）UC 浏览器漏洞与山寨客户端盗取信息

2014 年 5 月 11 日和 12 日两天，互联网用户遭到信息泄露的双重打击。5 月 11 日 UC 浏览器出现用户敏感数据泄露漏洞，只要在浏览器上搜索登录新浪微博等社交网站，登录信息就有可能泄露给黑客等不法分子。5 月 12 日出现了山寨的网银客户端和微信客户端，这些仿冒的客户端通过在手机软件中嵌入钓鱼网站，骗取网民提供身份证号、银行卡号等重要信息窃取用户资料。

（11）苹果公司承认 iPhone 存在安全漏洞

苹果公司承认 iOS 设备上存在一项默认激活且并未公开的服务。通过该服务能够在用户不知情的情况下，通过 WiFi 从其 iPhone 上窃取到通讯

① OpenSSL：OpenSSL 是一个强大的安全套接字层密码库，囊括主要的密码算法、常用的密钥和证书封装管理功能及 SSL 协议，并提供丰富的应用程序供测试或其他目的使用。

② Cookie：Cookie 是指某些网站为了辨别用户身份、进行会议跟踪而储存在用户本地终端上的数据（通常经过加密）。

录、短信、照片、地理位置等个人信息数据。

（12）1000 元竟能买到 1400 万条用户信息

由于网购的飞速发展，快递公司规模也逐渐庞大，其存储的用户信息也越来越多。在一起互联网网络信息案中，犯罪嫌疑人利用快递公司用户信息保存不严密的漏洞，窃取了 1400 万条快递用户信息并非法出售，牟利仅为 1000 余元。其之所以能够在公司数据库中畅通无阻，是因为快递公司的后台安全性很低，仅凭简单的技术手段便能轻易破解，前后窃取过程仅用 20 秒。

在正常情况下，快递公司的用户信息是需要被销毁的，而且还需向管理部门报备，而这家快递公司却保管着包括收货和发货双方的姓名、电话号码、住址等个人隐私信息。

（13）摩根大通银行信息泄露牵连 7600 万个家庭

2014 年 10 月 3 日，摩根大通公司公布电脑发生数据泄露问题，泄露信息包括用户姓名、地址、电话号码、邮箱等，除此之外，与用户相关的银行内部信息也遭到泄露，泄露的数据更是牵连到 7600 万个家庭和 700 万家小企业。

用户是金融企业的根本。用户信息的安全防护关系到企业的生存发展，一旦大面积用户信息遭窃，影响的不仅仅是企业的利益链，更为严重的是多年积累的用户信任感可能会消失殆尽。

向前追溯，最经典的案例莫过于将信息安全问题推入公众视野的“Wikileaks”（维基泄密）事件。

2010 年 3 月 15 日，美国军方反谍报机构 2008 年制作的的军方机密报告被泄露到网上，报告中提到了 Wikileaks 网站，直指其行为对美国军方机构的“情报安全和运作安全”构成了严重威胁。在这份报告被泄露后，Wikileaks 又发布了 2007 年美国士兵在巴格达滥杀平民的视频片段。随后，在 7 月 25 日，该网站又通过英国《卫报》、德国《明镜》和美国《纽约时报》同时公布了 92000 份阿富汗战争中美军的机密文件。机密文件中指出，驻阿富汗美军有许多不为人知的秘密，其中包括很多误杀平民事件，误杀总数达数百名，这些无辜平民多数被误认为是自杀式炸弹的袭击者而遭到射杀，而其真实身份往往只是司机或摩托车手。除此之外，阿富汗平民的

生命随时受到北约联军“373 特遣部队”的威胁，这支部队的职责是击毙或者俘获塔利班头目，同时，他们在行动中无需对塔利班头目进行审判。因此，在其执行任务的过程中，为了完成击杀任务，“373 特遣部队”会随意以牺牲无辜平民的性命为代价，无论是妇女、儿童，甚至是警务人员也不能幸免。

由于当时奥巴马的阿富汗战略正处在瓶颈期，而这些泄密文件可能会使民众对阿富汗战争的质疑不断加剧，这无疑是火上浇油。对此，五角大楼新闻发言人莫瑞尔表示美军军方将在泄密者绳之以法之前时刻处于戒备状态，以防更多信息泄露，但对于被泄露的机密文件内容缄口不言，称自己还无法确认数量如此庞大的信息。而 Wikileaks 总编辑朱利安·阿桑奇（Julian Assange）却在 8 月 18 日声称美国国防部愿意与其商议帮助 Wikileaks 审查其尚未发布的 15000 份阿富汗战争机密文件，以便屏蔽美军“线人”资料及其他敏感信息。

除此之外，Wikileaks 网站发布的泄密文件还显示，伊朗与基地组织和其他逊尼派极端分子存在深入的合作。例如文件中有资料细致地阐述了伊朗与塔利班和基地组织的联系细节。有美国官员称，美国多年来便一直收到伊朗向塔利班走私武器的报告，而在泄密文件中则直接指出伊朗官员与塔利班方面高层存在直接接触。而伊朗并不是走私武器的源头，其只是基地组织与提供武器国家间的中间人。

2010 年 10 月 23 日，阿桑奇对美国有线新闻网说，Wikileaks 公布的 40 万份伊战文件细致地记录了美军在 6 年的伊拉克战争中的作战情况，包括美军在战争中的见闻及所做的事情。文件中记录到，6 年来共有 10.9 万人在伊战中丧命，其中有高达 63% 的死者是伊拉克平民；对比 9 年丧生 2 万人的阿富汗战争，其死者人数为阿富汗战争的 5 倍，可以说是一场名副其实的“大屠杀”。文件中还记载了美驻伊部队虐待囚徒滥杀平民的事件，卡塔尔半岛电视台报道称，正是因为美军并未对该事件进行合理调查，由此受到了阿桑奇的抨击。2010 年 11 月 22 日，Wikileaks 网站表示，将在下一次大规模泄密行动中公开 300 多万份机密文件。对比上月公布的 40 万份伊拉克文件，此次的文件数量是对伊文件的 7 倍还多。而美国当局 27 日晚表示，拒绝就该拟泄露事件与 Wikileaks 网站进行谈判，并宣称

Wikileaks 持有密件的行为违反了美国法律。

而在美国当局宣布拒绝谈判的第二天，Wikileaks 网站公布了美国驻外使馆给美国国务院发送的 25 万份秘密文件电报，同时美国《纽约时报》、英国《卫报》等媒体对电报内容进行了报道，其中还有部分有关中国的内容。《纽约时报》报道的一份电报称，有“线人”向美国大使馆报告说中国政府招募黑客攻击并侵入了谷歌在中国的电脑系统，还称自 2002 年以来，这些被中国聘用的黑客已经成功入侵了美国政府、西方盟国、达赖喇嘛以及美国商业企业的电脑系统。

2. 大数据安全应对策略

大数据将成为推动世界经济发展的重要力量。但由于网络环境的高度开放性及相关法律的缺位，使得大数据时代下的互联网环境变得极度复杂。软件系统漏洞时有发生，应用程序使用人员复杂，社会缺乏对网络安全的重视，带来了一系列的大数据安全问题，其中数据安全与隐私保护问题尤为突出。

大数据为互联网安全带来挑战的同时，其自身的特性又使其能够在网络安全防护方面发挥作用。大数据处理技术的思想可以指导网络安全的发展方向。例如，可以对互联网数据进行收集整合，通过数据分析，构架出网络安全体系，找出安全问题，从而提醒有关部门采取相应的防护措施。如今已经有企业采用安全基线与大数据分析技术实现了网络异常行为及安全威胁检测的功能。2013 年，某研究机构发布的安全简报指出，大数据分析技术能够给信息安全领域带来诸多产品，这些产品将给整个安全领域带来重大转折，例如用户身份认证、信息安全事件管理、网络监控等等。简报还预测 3 年后会有大批金融、国防等行业的企业对网络数据进行分析，目的是找出网络安全潜在的威胁，而扫描的网络数据量至少会达到 10TB。

既然大数据能够应用到数据安全防护中，那么就能根据这个特性制定相应的大数据安全策略。由于策略的制定离不开对各领域内大数据安全需求的把控，因此只有了解各方面的需求，才能更好地制定大数据安全策略。

（1）互联网行业

大数据的应用与用户信息的安全问题在互联网行业中是最经常遇到

的，也是最受关注的。互联网的攻击行为具有隐秘性、复杂性以及难以界定性。首先，网上行为的增多往往会使相关的互联网企业受到更多难以察觉的攻击，而这些攻击造成的后果又仅仅是使服务器崩溃或停止工作。更令人担忧的是，面对如此隐蔽的攻击方式，相关企业很难保护数据完好不被侵害。其次，由于互联网的关系复杂性，公众个人隐私信息和企业机密信息会涉及诸多的技术领域。除此之外，由于很少有专家能够同时具有完备的法律和专业技术知识，这使界定损失类别及侵权主体变得十分困难。例如，批量用户数据的泄露，损失是属个人隐私还是商业机密传播，侵权主体是个人还是企业，这些都是需要界定清晰的问题。因此，互联网企业的大数据安全需求是：数据安全储存，加强安全分析频率与监管力度，制定专业性的法律法规和行业规范。

（2）电信行业

在大数据时代下，电信运营商需要具备很多能力。大量数据的产生需要运营商妥善处理好数据保密、用户隐私与对外合作等问题。首先，运营商需要明确数据价值并对其进行合理分类。其次，由于数据往往存在于各个系统中，来源庞杂，因此需要运营商具有高效收集整合信息的能力，并在收集过程中要保障数据完好且保密。在与其他企业合作时，运营商需要具备业务与数据相转化的能力，同时在建立数据开放访问机制时能够有效保护用户隐私与企业核心数据，防止其被不法分子窃取利用，给公司带来损失。因此，电信运营商的需求是：保障重要数据信息的安全与完整，维护用户利益与个人信息，同时实现数据价值最大化。

（3）金融行业

在金融行业中，金融系统中的个体相互联系，这导致了其安全风险也多方位地存在。而金融行业对于信息的可信性、保密性、安全性、稳定性和可用性要求很高。在应对复杂应用时，系统还要具有高速处理及备份数据、完善的管理能力及足够的灵活性的特点。但是由于金融业务链条增长、云计算的普遍应用等使得系统的内部复杂度增加，加上数据处理不当等原因，都使已经在数据安全方面追加投资和技术研发的金融行业并未过多地减少所面对的风险。

但是金融领域业务链条的拉长、云计算模式的普及、自身系统复杂

度的提升和使用数据方法的不正确，不仅抵消了在数据安全方面追加的投资，并未有效降低所面临的风险，而且大大增加了金融业使用大数据的潜在危险性。综上所述，金融行业要想安全地使用大数据必须满足如下要求：使用大数据安全及相关技术提高金融机构在各个环节（例如数据管理与访问控制、算法、网络等方面）对内的监管控制能力，从而提高服务水平，将隐患和风险降低到可接受的水平。

（4）医疗卫生行业

医疗数据的暴增加重了相关数据存储的压力。医院业务的连续性一定程度上取决于数据存储是否安全可靠。如果系统发生故障，数据损失首当其冲。只有快速恢复数据至检查点才能将医院受到的损失降到最低。此外，与医疗相关的数据敏感性很强，以至于极少医院主动“捐献”数据给其他组织或个人。同时，相关技术手段的局限性也大大降低了数据利用率。因此医疗行业需要大数据安全技术满足安全性和机密性，数据隐私性更加重要。此外还需要有安全且完备的数据存储、管理和备份，这样就能从多方面（诊断、管理、决策甚至药品的开发）提高医院的工作效率，从而更好地服务于患者，增强医院的竞争力。

安全的重要性已经不言而喻，所以恰当的应对策略就更为重要了。

（1）数据完整的防护

在考虑大数据发展的同时必须防止数据的丢失。安全问题在信息时代越来越多，对加密技术的灵活性和针对性的要求也越来越高。因此多模透明加密技术就成为最佳选项。这种技术结合了对称和非对称算法的优点，在不损失加密质量的同时更加灵活。处理方式越灵活，越有利于为大规模的数据安全提供保障。此外，在透明加密技术的帮助下，人们几乎感觉不到大数据的加密。该技术是基于系统内核的，这意味着它将具有更好的兼容性。既然我们要保护大数据安全，那么数据本身就应该是我们考虑的起点，因此我们最好使用加密软件。针对性强、防护全面的加密软件像哨兵一样保护了大数据的发展。对于企业来说，为了防止数据丢失，拥有快速检测数据威胁的能力是非常重要的，目前部分企业已经能够做到这一点。

（2）大数据不同于关系型数据库

大数据和关系型数据库，这二者看似差别甚微，实际上有很大的区

别。首先，它们具有不同的实时性，数据量也有差别。其次，它们的分布式架构也不尽相同，而分布式架构正是给安全防护带来独特困难的“元凶”。此外，大数据在存储与查询时采取与后者不同的模式，此外还需要协调不同网络会话。在大数据环境中，安全产品中有很多技术已经处于失效状态，其中包括监视与分析日志、发现数据以及评估漏洞等方面。因此，需要在架构层面上重新设计安全工具，以满足大数据环境中的安全需要。

（3）大数据加固网络层的安全策略

进行数据安全开发时，将数据结构化是一个好方法。该方法降低了数据处理和分类的难度，同时也方便了数据管理和加密。这样当发生非法入侵时，系统就可以准确高效地分辨出入侵行为，从而保证了大量数据在使用前不会被破坏。这种方法提高了系统的效率，但本质上并没有改变数据安全格局。数据结构化已经成为安全模式的发展趋势。

作为当前数据安全模式的常规做法，分层构建需要进一步完善。同时随着网络攻击次数的暴增及云计算造成的攻击方法隐秘性的增强，现有的端点安全模式已暴露出明显的弱点，因而使网络层受到强大的压力。所以我们应该在维护端点数据安全时重点考虑网络层。这要求我们在把数据结构化、辨识智能化与本地系统的监控机制结合起来时，只允许常态数据运行。

（4）本地数据层面的安全策略

在大数据时代，数据可以带来丰厚的经济收益，这也诱发了许多信息泄露事件，其中很大一部分来自内部。因此，对端点而言，本地安全防护系统看上去完整而成熟了，但实际上相差很大。这就要求调整安全防护思路，在本地安全策略中加入内部监控功能。为防止人为故意破坏，应使用纯数据模式。此外还应重视加强各环节的协作。在处理数据时数据调用有很大的风险，要想避免这种风险就要进一步划分链接，改进存储及缓存方式。

数据存储作为“终端”，受到了高度的重视，但其安全保护措施仍然需要加强，这样才能与新的数据模式相适应。这要求完善数据逻辑策略，作用于存储隔离与调用之间。

在大数据领域，只有少数开发资源被投入到增加安全功能中，而其他功能，例如分析功能、易用性和可升性，占据了大部分资源。此外还有一个显著的问题：大多数系统缺乏配套安全产品，而即便是有，也难以应对

常见威胁，而且非关系型数据库、Hadoop[①] 等无法包含大多数安全产品，因此用户自己构建安全策略就极其重要。本地安全策略可能存在许多未知隐患，这就需要用户一边开发，一边完善自有系统[②]。

（5）个人层面的数据安全

对于个人用户来说，将数据存放在对方服务器中就意味着一种抵押，由于对方想取用时无须任何申请，用户对此束手无策，因此也谈不上什么保护隐私。对此有几点建议：

①采用匿名 IP 地址。禁止网站搜集和跟踪 Cookies，不使用不支持 Do Not Track[③] 请求的浏览器。

②加密数据。主要针对企业级用户，对于个人用户来说，当其将一个私密文件上传到网络上，最好在压缩时设置加密密码，这无疑让用户的数据多了一道屏障。

③拒绝不合理的权限要求。这主要是针对手机用户，现在的手机应用程序，尤其是部分国产软件不顾用户的实际需求，所要求的权限超出了其本身的功能范围。此外，垃圾软件在后台运行占用硬件资源，严重影响手机性能及用户体验。

④浏览网页时使用 HTTPS 协议。HTTPS 协议是可进行加密传输、身份认证的网络协议，比 HTTP 协议安全，这样就增强了电脑与服务器之间收发的信息传输安全性。

3. 大数据安全应用

大数据应用，是利用大数据分析的结果，为用户提供辅助决策，发掘潜在价值的过程。本节主要探讨大数据在安全方面的应用。

---

① Hadoop：Hadoop 是一个由 Apache 软件基金会开发的分布式系统基础架构。用户可以在不了解分布式底层细节的情况下，开发分布式程序，充分利用集群的威力进行高速运算和存储。

② 孟威 . 大数据下的国家网络安全战略博弈［EB/OL］.［2014-08-11］http：//cpc.people.com.cn/n/2014/0811/c68742-25444087.html

③ Do not track：各大公司包括谷歌、微软的网络浏览器中添加一个“Do not track”（不追踪）按钮，让网民能够控制自己的隐私信息被追踪的情况，以此作为解决网络隐私问题的方案之一。