# 基于计算机大数据的信息安全处理技术研究

任丙权

（北京英菲尼迪科技有限公司，北京 100089）

**摘　要**　计算机大数据是以容量大、类型多、存取速度快、价值密度低为主要特征的数据集合。在计算机、互联网、云计算等高端技术迅猛发展的背景下，社会产业类型呈现出多样化特点，大数据技术也在社会各领域被普遍推广和应用。但是，受到人为操作失误、计算机病毒以及网络黑客等因素的影响，用户的计算机数据库也面临着较为严重的安全风险隐患，为了提高各种数据信息的利用价值，给社会各领域正常的生产与经济活动提供更加精准、确凿的数据支持，给计算机用户提供一个安全稳定的办公生产环境，广大技术人员应当有效运用信息安全处理技术，为计算机大数据构建一道坚固的安全防护墙。基于此，本文将着重围绕计算机大数据的信息安全所面临的问题以及信息安全处理技术要点展开全面论述。

**关键词**　计算机　大数据　信息安全　数据库

**中图分类号**：TP393.08　　**文献标识码**：A　　**文章编号**：1007-0745(2022)08-0016-03

大数据时代背景下，社会各行各业产生的海量化数据被统一存在计算机数据库当中，但是，由于计算机操作系统通常会处于网络环境之中，在这种情况下，数据库内的各种数据信息极易出现丢失或者被盗现象，这不仅给广大用户造成巨大的损失，同时，计算机大数据系统也将面临瘫痪的风险。为了切实解决这一问题，消除影响大数据系统信息安全的风险要素，近年来，云计算、数据备份、身份识别以及防火墙等安全处理技术逐步在计算机大数据信息安全防护领域得到广泛应用，进而使大数据的信息安全风险降到了最低点，各种数据信息的利用价值也显著提升。

## 1 计算机大数据概述

大数据也称之为巨量资料，主要是对一些庞大的、具有重要意义与价值的数据进行专业化处理的一种应用型技术，与大数据技术相关联的学科包括计算机、数学、金融以及统计学，因此可以说，大数据是整合了多学科的一种为计算机用户及社会各领域提供强大数据支撑的高端技术。该技术具有容量大、类型多、存取速度快、价值密度低的显著特征。

其中，数据库的最大容量能够达到 KB 级，下一个层次级别分别是 MB、GB、TB、PB、EB、ZB、YB、NB、DB 级，据统计数据表明，目前，非结构化数据比结构化数据增长速度快 10 到 50 倍，因此，非结构化数据库是传统数据库容量的 10 倍到 50 倍。所谓非结构化数据是数据结构不规则或者不完整，没有预定义的数据模型，常见的非结构化数据包括各种格式的办公文档、文本、图片、XML、HTML、各类报表及视频图像等，而结构化数据是指存储在数据库当中行数据，这种数据可以用二维表结构来表示。数据类型多是指大数据具有很多不同形式的数据类型，比如文本数据、图像数据、视频数据、机器数据等，除此之外，还包括互联网的网络日志、视频、图片、地理位置等。存取速度快是应用大数据技术可以对海量数据进行实时分析，时间计算单位通常以 1s 为临界点，在处理各种数据信息时，通常会在 1s 之内形成答案。而价值密度低主要是指在海量的数据当中，挖掘出具有重要参考价值的珍贵信息，因此，价值密度低也是大数据的一个典型特征[1]。

## 2 基于计算机大数据的信息安全面临的问题

### 2.1 计算机病毒侵蚀

由于计算机系统往往接入互联网平台，这就使得计算机系统成为一个开放的共享环境状态，在这种状态之下，计算机病毒可以乘虚而入，直接破坏支撑数据库运行的计算机硬盘及软盘。比如 2007 年暴发的灰鸽子病毒，该病毒将自身伪装成计算机上的正常文件，而且能够有效躲避网络防火墙软件的监控，使其很难被用户发现。另外，黑客可以利用控制端地受感染的计算机进行远程控制，并进行多种危险操作，包括查看计算机用户的各种系统信息，同时，从互联网上面下载任意的指定恶意文件，来盗取用户的个人隐私信息。比如用户的个人网上银行账号、密码以及 QQ 号等。如果用户使用的计算机安装了摄像头，黑客还会随时

捕捉摄像头拍摄下来的所有画面内容，这样既给计算机系统的正常操作带来不利影响，而且计算机数据库中的用户信息也面临着丢失的风险。因此，计算机病毒对计算机系统的侵蚀，是数据库所面临的最大安全风险。

### 2.2 网络数据的不安定因素

随着互联网时代的到来，网络空间当中存在大量的数据，这些数据分布范围广，涉及类型多，比如传感器、社交网络、网络日志、电子邮件、记录文档等，随着数据量的不断攀升，计算机用户个人隐私信息也面临着严重的安全风险。由于数据库具有海量化特点，因此，数据库当中存储的数据还包含了大量的个人隐私信息、客户企业、企业运营数据等，一旦计算机系统接入互联网平台，那么，这些数据的被盗风险也将显著提升。目前，受到网络不安定因素的影响，个人隐私受到侵犯的表现特征如下：第一，数据存储阶段面临信息被盗风险，在这种情况下，用户无法对数据的收集与存储进行有效控制，一旦出现这一状况，用户数据丢失的可能性将大幅增加。第二，数据传输过程中面临信息丢失风险，在网络云环境下，各类数据的传输环境将变得更加开放，如果采用传统的物理隔离法对数据进行防护，其防护距离也受到严重限制，尤其在电磁泄漏与非法窃听等安全问题频发的背景之下，物理隔离法已经无法保障用户数据信息的安全性。第三，数据处理阶段所面临的安全风险，由于云服务所提供的技术属于虚拟技术序列，这种技术的基础防御功能较差，尤其在处理大批量数据时，需要完整的访问控制与身份认证管理，这就给云服务动态模型管理增加难度，在这种情况下，极易出现密钥丢失、认证失效、黑客攻击、伪装账户身份等现象的发生，进而给用户的个人数据信息埋下了重大的安全风险隐患。由此可以看出，计算机用户的个人数据信息在网络环境下所面临的安全风险类型呈现出多样化特点[2]。

### 2.3 网络黑客攻击

当网络空间中的数据呈指数级增长时，大量数据的堆积给网络黑客留下了可乘之机，在这种情况下，网络黑客将直接通过计算机系统漏洞攻入数据库，数据量越大，被网络黑客发现的可能性越大。一旦网络黑客攻入计算机系统以后，将冲破一些没有真正意义与价值的数据，而直接获取对企业或者个人具有珍贵价值的数据信息，这时，计算机用户个人数据信息的泄漏风险也将不断攀升。比如网络黑客将借助于社交网络、电子邮件、微博、电子商务、电话等信息直接向用户的计算机数据库发起攻击，并且可以对海量化的数据资源进行操控，一旦出现这种情况，用户的主动防御功效也将丧失。

### 2.4 计算机用户的主观人为失误

用户在使用计算机过程中，由于疏于防范，以至于一些外来人员可以随意登录计算机系统，尤其是外来人员利用 U 盘、光盘、移动硬盘等外部存储介质拷贝计算机内部资料时，事前并未对这些外部存储介质进行杀毒处理，使得存储介质当中携带的病毒与木马程序直接攻入用户的计算机系统当中，在这种情况下，用户计算机数据库中大量有价值的数据信息将面临严峻的安全风险。比如用户未及时对杀毒软件进行升级，或者未及时安装下载漏洞补丁，使一些木马程序直接从计算机系统漏洞当中进入到数据库，这时，用户所使用的计算机也将面临瘫痪或者中止运行的风险。

## 3 基于计算机大数据的信息安全处理技术

### 3.1 云计算技术

云计算技术是保障计算机大数据信息安全的一项关键技术，该技术在保护大数据信息安全时常常通过分布式与并行式两种形式进行：一方面，利用云计算技术能够大幅提升数据的计算精准度与数据的分析处理速度；另一方面，云计算技术可以对各类数据信息进行优化配置与组合，进而来突显数据信息的实际利用价值。目前，基于云计算技术的云安全基础服务主要包括云用户身份管理服务、云访问控制服务、云审计服务以及云密码服务。其中，用户身份管理服务涉及用户身份的提供、注销与认证，在云环境下，能够实现身份联合与单点登录，但是，在用户登录过程中，需要确保用户数字身份隐私性，这样才能保障用户个人数据信息的安全性。云访问控制服务，主要是将基于角色访问控制、基于属性访问控制以及自主访问控制等模型移入云环境当中，这时，各种资源服务的兼容性与组合性能够突显出来，正是这种组合式的授权模式，才使计算机大数据的安全防护屏障变得更加坚固。云审计服务主要是由第三方对云环境下用户数据信息提供审计服务，在审计过程中，用户个人数据信息不会被披露，因此，这种云审计服务具有合法性与合规性。而云密码服务主要是为了满足用户加密与解密的需求而设置的一种云服务模式，这种模式大大简化了用户的登录密码，使得用户的数据信息更易于管理和有效控制[3]。

### 3.2 数据备份技术

数据备份技术主要是为了防止计算机系统文件丢失，而将所有数据或部分数据从主机硬盘复制到其它

存储介质当中，一旦计算机系统瘫痪，用户可以直接在备份当中寻找相关数据信息。但是，对于有些备份来说，处于无法恢复状态，即便已经做了备份处理，但是却无法找到源程序，这种备份方式很显然是没有任何意义的。因此，用户在备份数据信息时，应当额外进行备份，以防止数据无法得到恢复。目前，数据备份的基本操作主要分为四种方式：全量备份、增量备份、差量备份与立即备份。

其中，全量备份是将计算机系统中的全部内容进行备份处理，比如备份对象是磁盘卷，那么，在备份过程中，用户应将磁盘卷的全部内容拷贝到其它存储介质当中，这种备份方式便于全量恢复，但是，全量备份却容易造成资源浪费，比如一些冗余的数据信息也将被一同备份到存储介质，这就增加了介质的存储负担。增量备份是在全量备份的基础上，对计算机系统新产生的数据信息进行更新与备份，这种方式的备份耗时短，但是，在恢复备份时，需要执行最近一次的全量备份，否则难以提取出增量备份的内容。差量备份是将所有新产生的数据信息进行备份，这些数据都是最近一次全量备份以后所产生的，这种备份方式与增量备份最大的区别在于差量备份是以全量备份为检查点滴，而增量备份则是以上一次的全量备份为检查点。立即备份主要是对即时发生的数据信息进行备份，这种备份具有很强的随机性，往往需要借助于人工操作来完成，因此，在备份时，用户可以自行选取备份时间。利用数据备份技术可以有效防止数据信息的丢失，同时，也使数据信息的利用价值得到切实体现。

### 3.3 身份识别技术

身份识别技术是一种严谨、高效、实用的信息安全处理技术，究其原因主要是由于每一个人只有一重身份，这重身份无法替代，无法更换，无法伪造，而且，在实际应用过程中，身份识别技术所需要的各类信息便于采集，便于处理。比如以密钥验证技术为例，这种技术主要包括对称密钥与非对称密钥，对称密钥主要是加密密钥能够从解密密钥当中推算出来，这些算法也叫秘密密钥算法或单密钥算法。对称密钥更加依赖于密钥，一旦密钥泄漏，也就意味着任何人都可以数据信息进行加密与解密。而非对称密钥则是指一个加密算法的加密密钥，它与解密密钥是不同的，因此，由一个密钥将无法推导出另一个密钥。通过密钥技术能够大幅提升用户数据的安全性能。另外，计算机用户最为常用的一种身份识别技术是口令识别法，即服务器需要采用用户名与口令对用户进行认证，同时，还需要提供口令更改工具，以便于口令泄漏以后及时对其进行修改。在选择口令时，一般遵循“易牢记、不易猜、不易分析”的原则，这样才能大幅降低用户数据信息的安全风险[4]。

### 3.4 防火墙技术

防火墙技术是计算机大数据信息安全领域最为常用一种安全防护技术，该技术主要包括数据包过滤技术、代理技术以及 IP 地址翻译技术。其中，数据包过滤技术主要是根据一定的过滤规则对用户的数据包进行针对性的过滤筛选。用户在建立局域网络时，在跳转路由上面添加数据包过滤规则。比如访问 IP1 的数据报文将会被跳转至相对应的路由器，访问 IP2 的数据报文会被丢弃，当过滤规则确定以后，路由器则分别对这些数据进行过滤与筛选，进而形成白名单与黑名单，顾名思义，属于黑名单的数据将受到访问限制。代理技术则是在用户的被保护数据的外层添加访问代理的方式，来保护数据安全，当网络资源试图访问用户数据信息时，防火墙将直接把请求信息发送给代理服务器，这时，代理服务器根据被访问数据的性质而采取保护隔离动作。而 IP 地址翻译技术则是将用户的 IP 地址资源进行分类，然后在局域网当中，通过配置与之相对应的掩码与网关来设定 IP，当用户具有独立的 IP 地址以后，用户的数据安全也将得到可靠保障，可见，防火墙技术在保障计算机大数据的信息安全方面扮演着重要角色。

## 4 结语

随着计算机大数据信息安全防御技术的日渐成熟，技术类型也逐步呈现出多样化特点，因此，为了给计算机大数据构筑一道坚固的防护墙，计算机用户应当不断对安全处理技术进行创新和改进，并积极应用一些新技术来增强和改善计算机系统的安全性能，进而使计算机大数据的实际应用价值逐步突显出来。

**参考文献：**

[1] 陈荣 . 基于计算机大数据的信息安全处理技术 [J]. 中国新通信 ,2021,23(21):136-137.
[2] 何振贤 . 基于计算机大数据的信息安全处理技术分析 [J]. 中国管理信息化 ,2021,24(07):167-169.
[3] 陈保 . 大数据背景下计算机信息安全处理技术探究 [J]. 南方农机 ,2019,50(06):169.
[4] 薛佳桦 . 试论基于大数据环境下的计算机信息安全技术 [J]. 电子制作 ,2018(12):53-55,46.