文章编号：2096-1472(2018)-07-50-03
DOI:10.19644/j.cnki.issn2096-1472.2018.07.014

基于大数据分析技术的网络安全系统软件开发与设计

王春莲
(德州职业技术学院计算机信息技术工程系，山东 德州 253000)

摘　要：网络环境下软件运行受到病毒威胁，存在安全隐患，如何通过软件设计来提升病毒防御能力，是当前开发中重点研究的内容。本文以此为出发点，结合大数据分析环境下软件的网络安全威胁特征，重点分析软件开发与设计中逻辑组成特点及使用安全需求，并提出通过数据跟踪定位来降低病毒入侵风险的技术措施，其目的在于提升网络环境下的软件安全性，在病毒黑客入侵后也能最大程度保障内部数据安全。本研究中所论述的技术观点，在实际开发应用中还需要结合运行环境深度优化。

关键词：大数据；网络安全；软件开发
中图分类号：TP311　**文献标识码**：A

Development and Design of the Network Security System Based on Big Data Analysis

WANG Chunlian
(*Dezhou Vocational and Technical College*,*Dezhou* 253000,*China*)

Abstract: Under the environment of large data analysis,the function of network security system is seriously affected by the operating environment.How to improve the ability of virus defense through software development and design is also the key research content in the current network system development.This paper takes this as the starting point and combines the characteristics of the network security system under the large data analysis environment,and focuses on the analysis of the logic components and the use of security requirements in the software development and design.This paper proposes a technical method to reduce the risk of virus invasion through data tracking and location.The purpose is to improve the compatibility and security of software in large data analysis environment,and to ensure the maximum security of internal data after the virus hacker invades.The technical points discussed in this paper should be further optimized in combination with the operation environment in the actual development and application.

Keywords:big data analysis;network security;software development

1 引言(Introduction)

基于大数据分析环境下进行网络安全软件开发设计，首先需要对逻辑进行分层，确定软件功能实现的逻辑组成，逻辑分层需要从信息传输信息提取与网络端口对接三个层面进行[1]。通过建立起逻辑数据库来实现信息分类存储、数据提取与网络端口对接。并对不同功能实现所需要的数据进行隔离，避免在数据传输过程中产生干扰。逻辑分层后不同数据库所对应的功能也会处于隔离状态，在软件使用过程中通过网络平台来下载海量数据，需要在短时间内对数据的安全性进行检验。当初检验得到的结果显示逻辑层内存在信息精准度不足的问题，则会进行风险来源判断，以免数据信息使用过程中出现影响安全性的问题。控制管理过程中数据库分层可以针对所接收到的功能请求自动对接，通过逻辑分层方法，帮助最大程度保障软件运行环境安全性，逻辑分层的最终环节是数据库分层，依照数学建模方法来对数据库对应区域进行划分[2]。大数据环境下软件接收到数据传输请求后，会自动启动所对应的数据库，通过这种方法最大程度确保软件运行过程中逻辑处理的效率。

2 大数据分析下基于网络安全系统的软件逻辑组成(Software logic based on network security system in big data analysis)

2.1 海量数据提取

海量数据提取则是面对大数据分析环境下文件处理需求较大的情况来进行，网络环境下软件运行需要面对大量数据信息处理任务来进行，为避免影响到数据抽取速度需要对数据安全性进行模糊判断[3]。模糊处理方法应对海量信息数据抽取具有明显的效率提升效果，海量信息处理，并不表示在安全结果上会受到影响，而是通过一种预分析方法，初步判断数据是否存在风险，并对潜藏风险数据进行划分进入到细致分析环节中。可以对安全数据与风险数据进行二次深入分

www.rjgcz.com

析，最终得到的数据抽取结果不仅能够满足软件使用中的逻辑需求，更可以避免风险数据影响到软件安全使用，海量信息处理后数据安全性会有所提升。大数据分析中的逻辑层结构如图1所示。

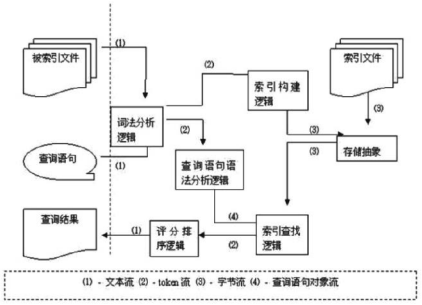

图1 大数据分析中的逻辑层结构

Fig.1 Logical layer structure in big data analysis

数据处理期间要以安全功能为前提，观察处于不同状态下，数据所遇到的风险威胁，通过信息预处理技术来降低风险威胁指数，为软件提供更安全的参数。海量数据提取中，所确定的算法要确保兼容性，观察各项控制方案中存在的功能隐患，通过预处理方法来提升功能性。

2.2 数据跟踪定位

数据跟踪定位主要是针对风险信息来进行，大数据分析环境下实现网络安全目标，需要软件程序开发中引入数据跟踪定位功能能够根据所分析得到的数据安全性结果，自动寻找数据来源[4]。在系统开发设计中也会对逻辑层进行数据库深度开发，当使用过程中产生风险数据参与分析的情况，通过数据跟踪定位，可以准确判断网络安全性受到的影响，并了解处于数据跟踪环境中造成风险的主要因素，整理成为经验文档，数据使用过程中再次出现此类问题也能直接启动风险控制数据库。面对大数据分析环境下海量信息处理任务，仅仅依靠数据来源逐一划分不仅浪费时间，最终功能实现也会受到影响。数据跟踪定位功能引入，将数据处理与分析任务所针对范围进行缩减，仅仅针对有效数据来进行，提升速度的同时也能避免非法请求通过软件控制层。数据处理过程中会自动形成软件控制方案，对网络平台下潜藏的风险进行规避，如何利用数据处理来实现更高效的安全防控，则需要借助数据跟踪定位功能来完善。

3 大数据分析下基于网络安全系统的软件设计 (Software design based on network security system under big data analysis)

3.1 文件系统的节点选择

大数据分析环境下，基于网络安全系统对软件进行开发设计，首先需要对文本系统的参数运算节点进行选择，按照数学排列特征方法来选择文件提取过程中的参数节点。但运行使用期间大数据分析环境下受网络安全性影响，仍然会对所选择的节点做出优化调整。网络设备、外部系统、原始流量分别采用sys-log、webservice、sftp和netflow进行收集，收集到的数据通过flume或ftpd进行传输进入到数据处理队列kafka中，对于收集到的文件直接保存到HBase中。节点选择需要从数据结点与管理信息节点两方面进行，并对所选择的节点进行网络串口连接，实现软件与网络平台的对接，数据信息提取中也能够通过网络窗口来高效传输。控制节点选择服务器构建图如图2所示。

节点设计是对数据安全性的有效识别方法，因此需要在安全控制范围内来开展，对于节点控制期间可能会产生的常规风险隐患问题，通过数据库相互联系便能够高效解决文件存储和远程数据读取，综合判断出信息安全控制范围[5]。当信息预算处理中产生影响安全性的问题，可以通过节点追踪来判断风险数据所处区域，尤其是对于文档所受到的病毒黑客攻击，可能会造成内部文件丢失问题。Node1—Node11、模拟日志采集器和模拟流量发送器均接入一台交换机组成局域网，IP地址为192.168.1.11—192.168.1.21，其中Node1—Node11组成大数据集群和分布式存储，Node1和Node2为主备管理节点，Node3—Node9为数据存储节点，Node10为大数据集群总线调度节点，Node11作为大数据分析平台前端展示节点选择需要形成整体性，相邻结点之间能够相互检测是否存在黑客病毒，从而发挥更高效的控制效果。

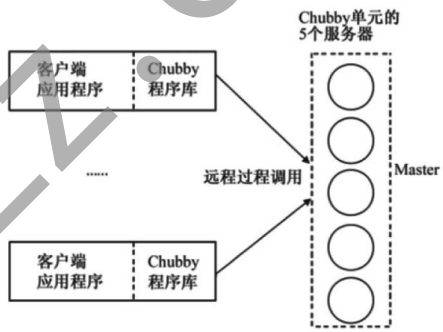

图2 控制节点选择服务器构建图

Fig.2 The control node selects the server build diagram

3.2 文件存储副本设计

文件存储的副本设计，是应对软件使用过程中内部文件受到病毒黑客攻击后，产生数据信息丢失的情况。副本与文件正本存在的本质差异性在于读取范围，副本仅仅可以对其中的内容进行读取并不具备篡改能力，因此在文件存储过程中对副本做出合理的设计可以避免病毒直接入侵原数据。对于副本的设计还应该考虑是否会影响到正式文件使用功能，副本与源文件在读取权限方面要做出区分，避免打开时产生干扰影响到文件读取速度。副本与源文件在保存路径方面也要有严格区分，如果存储过程中出现数据，存储路径相互干扰的情况，不利于文件中数据存储安全性提升。首先在Android的布局文件下中添加文件控件com.baidu.mapapi.map.MapView，创建文件的MainActivity继承Activity，管理文件生命周期，然后在的onCreate方法中获取文件控件引用，开发框图构建操作图如图3所示。

副本设计可以采用路径构建方法来实现，在文件保存过程中会自动生成副本，不同软件使用者对文件的访问权限也会做出选择，可以通过源文件与副本文件整合的模式来提升软件运行速度。文件存储后副本也会随之更新，在内容上与原文件保持一致。属于大数据分析网络环境下，文件公开的部分也主要针对副本进行，这样不同黑客病毒攻击情况也是

www.ircz.com

直接在副本范围内开展，源文件在大数据分析环境下性能保证绝对安全，并不会受到副本使用情况影响。

现在大数据环境下软件开发及设计中会针对安全性进行重点提升，但仍然不能避免网络环境中出现的突发情况，威胁到软件内部文档数据安全。对此副本设计中也会跟对风险进行维护，进行副本与源文件之间的对接控制。副本源文件形成整体后，既能保证数据在风险环境下相互隔离，同时也能确保大数据分析任务高效实现。

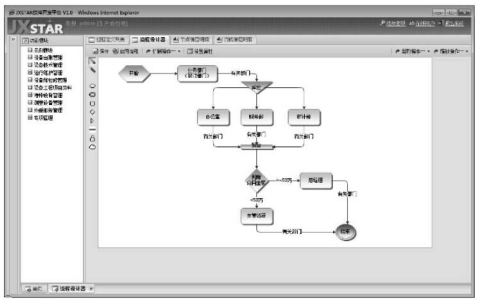

图3 开发框图构建操作图

Fig.3 The development block diagram builds the operation diagram

3.3 数据恢复系统设计

数据恢复系统可以帮助已经受到病毒攻击的软件恢复常规使用功能，软件内部文档中所存储的重要信息，一旦受到病毒攻击，便会出现信息丢失。数据恢复系统可以帮助恢复到软件初始化输入状态，也可以根据操作者使用需求来进行恢复状态设定，对于这部分系统的设计主要从预备数据库开发层面来进行。在软件正常使用中对重要信息的数据库进行备份，这样一旦受到病毒攻击，出现信息丢失或者文件损坏的问题，也可以通过数据恢复来达到原有状态。数据丢失后恢复系统会自动进行检索判断所丢失的具体情况，当软件系统受到严重的病毒入侵时，常规使用功能将难以实现，此时需要进行系统重新恢复。损坏严重的文件，通过恢复也难以打开，便会彻底删除。因此恢复系统中会涉及提醒功能，操作者在存储一些重要文档时会接受到备份提醒，避免出现风险软件恢复到初始化状态影响使用者文件安全。对于一些损坏不明显的系统文件，会进行局部恢复补偿，恢复后系统并不会进入到最初状态，也能最大程度恢复所存储的文件档案。

4 大数据分析环境下基于网络安全系统的软件开发实现(The realization of software development based on network security system in big data analysis environment)

4.1 脚本测试代码构建

脚本测试代码可以帮助快速判断大数据分析环境下提取的信息，是否符合安全规定。目前计算机软件开发通常基于Windows 10操作系统，该系统具有自我防护能力可以针对网络环境中的病毒自动查杀控制，因此在该环境中要考虑所开发软件功能与系统自身防御系否会产生冲突，要重点提升软件设计中的兼容性，对于脚本进行测试判断代码与系统中的程序文件是否存在冲突。以运行环境为依据对软件做出优化构建，系统使用中同时面对大量处理运算任务。脚本测试是判断其安全性最便捷的方法。对于脚本测试代码的构建需要通过网络环境中确定常规数据范围，软件投入使用后在网络环境下能够自动更新，因此脚本也会存在更新情况，如果将更新后的使用方向同样引入其中，所构建的代码运算也会增大消耗时间。针对基础数据构建代码测试框架，软件投入使用后更新下载是建立在原框架基础上所进行，因此只要掌握源文件技术特征，后续脚本代码安全性检测也可以在此基础上快速完成。

4.2 海量信息处理模型建立

信息数据处理模型建立，要考虑在大数据分析环境下要考虑海量信息，同时处理时对于系统资源的占用情况，数据处理模型建立重要考虑系统投入使用后的均衡性。病毒入侵是长期存在的，随着软件功能更新所面对的病毒及黑客入侵情况也会有所转变。信息数据安全性检测期间，可以模拟出网络环境下所存在的病毒攻击对系统进行内部攻击控制，从而判断软件系统中存在的风险部分，通过信息处理系统模型建立来确定需要继续强化的部分。软件模型建立后还需要对风险发生类型进行划分，要确保软件处于病毒攻击环境下也能最大程度保障内部文件安全，所选择的程序汇编语言与Windows 10操作系统之间，需要保持连贯性。

提升系统内部自我防御能力，检测到风险和黑客攻击后，也会自动向控制中心发出警报，软件是基于计算机网络环境下使用的，因此软件安全性与网络安全是共同存在的。两者之间既存在相同点，同时也存在差异性，要基于相同点构建框架基础，并对各自特征做出保留。

5 结论(Conclusion)

云计算和物联网的迅速发展，越来越多的个人和企业选择将自己的业务迁移到大规模的数据中心，以此来降低本地的硬件成本和系统维护费用。数据中心存储的数据量十分庞大，管理系统的复杂性较高。从存储设备级别上看，数据中心为了控制成本，大量采用廉价存储设备，致使数据极易因硬件设备故障而丢失。这些都对海量数据存储性能、可靠性等方面带来了挑战。云存储是解决海量数据存储的最有效手段。本文提出的一种海量大数据云存储系统设计，与网络安全系统相互结合，是应用系统需要实现的关键技术。

参考文献(References)

[1] 李锦华.基于大数据分析的海量信息软件系统设计与开发[J].软件工程,2017(11):54-56.

[2] 刘世民,朱继阳,高敏,等.基于大数据的分布式网络安全管理平台的设计[J].信息技术,2016,40(11):0011-0012.

[3] 王萍.基于大数据技术的网络异常行为分析监测系统[J].电子技术与软件工程,2017(24):172-173.

[4] 陈祖斌,谢铭,胡继军,等.基于大数据和可信计算的信息网络安全自防御系统:CN106209817A[P].2016:0015-0016.

[5] 钟伟.大数据驱动的飞行信息物理融合系统的分析与设计方法[D].广东工业大学,2016.

作者简介：

王春莲(1973-)，女，硕士，副教授.研究领域：物联网，云计算.

www.ipcz.com