



(12) 发明专利

(10) 授权公告号 CN 117828644 B

(45) 授权公告日 2024. 07. 02

(21) 申请号 202410015209.9

(22) 申请日 2024.01.05

(65) 同一申请的已公布的文献号

申请公布号 CN 117828644 A

(43) 申请公布日 2024.04.05

(73) 专利权人 江苏商贸职业学院

地址 226011 江苏省南通市崇川区江通路
48号

专利权人 南通智大信息技术有限公司

(72) 发明人 曹金华 瞿国亮 顾林强 顾俐俐
金飞(74) 专利代理机构 北京金蓄专利代理有限公司
11544

专利代理师 洪涛

(51) Int.Cl.

G06F 21/60 (2013.01)

G06F 21/62 (2013.01)

G06F 21/55 (2013.01)

G06F 16/35 (2019.01)

(56) 对比文件

CN 115942262 A, 2023.04.07

CN 117224972 A, 2023.12.15

审查员 胡学岭

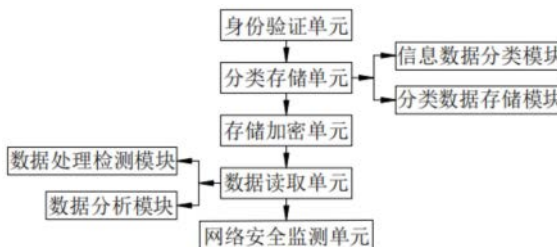
权利要求书4页 说明书11页 附图1页

(54) 发明名称

具有信息安全保护模式的计算机存储系统

(57) 摘要

本发明公开了具有信息安全保护模式的计算机存储系统,涉及计算机存储技术领域,为了解决信息数据存储时的安全性能下降的问题。本发明通过对数据进行去除重复数据、填充缺失值、数据标准化的处理可以最大程度的保证数据的完整性,对周期序列数据和非周期序列数据中的异常数据都进行提取,可以快速定位和识别这些异常数据,以便及时采取措施进行处理和纠正,根据加密可以将数据转换为密文,只有持有密钥的人才能将其还原成原始数据。即使数据在传输过程中被窃取,盗取方也无法读取到原始数据,从而保护了数据的安全性。



1. 具有信息安全保护模式的计算机存储系统,其特征在于,包括:
身份验证单元,用于:
进行系统操作时将操作者的管理身份进行身份认证,身份认证成功后操作者进行信息录入或者信息检索;
分类存储单元,用于:
基于身份验证单元中获取的录入信息,将录入信息根据关键词进行属性分类,属性分类完成后将每个属性的数据存入对应的存储器中,并将每个存储器中的数据进行数据备份;
存储加密单元,用于:
基于分类存储单元中获取的存储器中的数据,将存储器中的数据进行数据加密,并将加密完成的数据标注为加密数据;
数据读取单元,用于:
基于存储加密单元中获取的加密数据,将加密数据进行周期性检测,根据检测结果将加密数据中的异常数据进行确认,并将异常数据进行单独调取并显示;
网络安全监测单元,用于:
进行数据操作时对网络IP进行实时监测,若监测出异常IP时进行端口关闭,当网络IP进行正常信息操作时,进行实时网络流量监控分析,若通过分析网络流量数据有异常行为时,进行端口关闭;
所述存储加密单元,还用于:
将每个存储器中的录入信息数据进行加密段划分,并将每个划分完成的加密端进行加密序列建立;
根据加密端的数量,将加密序列作为矩阵的第一列,建立第一加密矩阵;
根据预设干扰因子对每一待加密段在第一加密矩阵对应行上的第一位置进行调节,调节完毕后对其进行第一加密;
基于干扰次数在预设密码集中获取对应的干扰密码,同时利用加密骰子掷出随机数;
在第一加密矩阵的每一行中标记与第一位置相隔随机数的第二位置;
将干扰密码依次输入到第二位置进行第二加密,建立第二加密矩阵;
当第二加密矩阵中的空白位置的第一数量大于非空白位置的第二数量时,确定当前加密程度不足;
依次剔除第二矩阵的最外围位置,生成若干个子加密矩阵;
分别计算每一子加密矩阵的加密秩,基于所有的加密秩生成秩密码;
在第二加密矩阵上标记剩余空白位置,并将标记后的第加密矩阵与同型矩阵进行叠加,获取标记位置产生的叠加数;
将叠加数输入到剩余空白位置上,获取全加密矩阵;
提取第二全加密矩阵中每一位置对应的数据,获取加密数据;
所述数据读取单元,包括:
数据处理检测模块,用于:
将存储器中的加密数据进行获取,将获取的加密数据进行数据处理;
其中,数据处理包括去除重复数据、填充缺失值、数据标准化;

数据处理完成后进行周期性检测,根据周期性检测结果判断加密数据是否为周期序列数据;

数据分析模块,用于:

当数据进行周期性检测后为周期序列数据时,对周期序列数据按照周期进行分割得到多组相同的第一子序列,判断第一子序列中的全部数值是否大于预设数值,若是,提取出第一子序列中大于预设数值的第一异常数值,并将第一异常数值作为第一异常数据;

当数据进行周期性检测后为非周期序列数据时,通过一维聚类方法对周期序列数据进行聚类运算,得到多个分割点,并利用多个分割点对周期序列数据进行划分,得到多组不同的第二子序列,从第二子序列中获取存在大于预设数值的第三子序列,并将第三子序列作为第二异常数据;

将第一异常数据和第二异常数据进行数据整合,整合完成后将整合数据发送至显示终端进行数据对应显示,其中,数据对应为将整合数据存在的存储器以及录入网络IP进行对应。

2.根据权利要求1所述的具有信息安全保护模式的计算机存储系统,其特征在于:所述身份验证单元,还用于:

操作者进行身份认证,身份认证为密码认证和账号认证,密码认证和账号认证后系统进行密码和账号的核对;

核对完成后对该操作者的职位进行识别,其中,不同的职位进行不同的系统操作;

身份认证成功后对操作者的网络IP进行验证,其中,将该操作者的网络IP与数据库中认证完成的网络IP进行核对,当操作者的网络IP与数据库中认证完成的网络IP一致时,该操作者能够进行下一步的信息数据操作。

3.根据权利要求2所述的具有信息安全保护模式的计算机存储系统,其特征在于:所述分类存储单元,包括:

信息数据分类模块,用于:

将录入的信息数据进行获取,并提取信息数据的关键信息片段;

根据关键信息片段的目标语义确认录入信息数据的信息类型;

其中,目标语义为将关键信息片段中位置连续的词语以文本最小单元进行最大长度的拼接,拼接后的得到拼接语义片段;

根据词性统计特征对拼接语义片段进行清洗,得到片段关键词,其中,词性统计特征从数据库中进行获取;

根据片段关键词确认录入信息数据的具体文本属性,并将具体文本数据作为该录入信息的信息类型。

4.根据权利要求3所述的具有信息安全保护模式的计算机存储系统,其特征在于:所述分类存储单元,还包括:

分类数据存储模块,用于:

基于信息数据分类模块中获取的录入信息数据的信息类型,将每个类型的录入信息数据进行长度确认;

将录入信息数据划分为若干个长度相等的段落;

根据段落的数量确认该录入信息数据的总长度,并根据录入信息数据的总长度计算出

该录入信息数据的大小；

根据录入信息数据的大小,将录入信息数据存储至对应的存储器中,其中,将存储器的剩余容量进行实时监控,录入信息数据的存量小于存储器的剩余容量；

将每个存储器中的信息数据进行数据备份,并且每个备份数据均进行唯一编码标号。

5.根据权利要求1所述的具有信息安全保护模式的计算机存储系统,其特征在于:所述分类存储单元,包括:分类单元,用于基于身份验证单元中获取的录入信息,将录入信息根据关键词进行属性分类；

所述分类单元,包括：

第一确定模块,用于：

将所述录入信息输入预先训练好的BERT模型中,输出录入信息包括的各个文本数据块的语义表示,根据语义表示结果确定文本数据块包括的关键词；

计算每个文本数据块包括的关键词与预设数据库中各个预设关键词的匹配度：

$$P_i = \frac{\sqrt{\left(\sum_{x=1}^N w_{i\partial_x}\right) \times \left(\sum_{y=1}^M \overline{w_{\epsilon_y}}\right)}}{\sqrt{\left(\sum_{x=1}^N w_{i\partial_x}\right) \times \left(\sum_{y=1}^M \overline{w_{\epsilon_y}}\right)} + \sum_{x=1}^N \sum_{y=1}^M w_{i\partial_x} \times \overline{w_{\epsilon_y}}}$$

其中, P_i 为每个文本数据块包括的关键词与预设数据库中第*i*个预设关键词的匹配度； ∂_x 为第*i*个预设关键词中第*x*个语义表示； ϵ_y 为文本数据块包括的关键词中第*y*个语义表示； $w_{i\partial_x}$ 为语义表示 ∂_x 在第*i*个预设关键词中出现的次数； $\overline{w_{\epsilon_y}}$ 为语义表示 ϵ_y 在文本数据块包括的关键词中出现的次数；*N*为第*i*个预设关键词中语义表示的数量；*M*为文本数据块包括的关键词中语义表示的数量；

根据最大匹配度对应的预设关键词确定每个文本数据库的数据类型,并进行聚类,得到若干个分类集合；

第二确定模块,用于：

分别确定每个分类集合中的聚类中心,确定各个聚类中心之间的距离；

在确定各个聚类中心之间的距离均大于预设距离阈值时,将若干个分类集合作为最终的属性分类结果；反之,重新将录入信息根据关键词进行属性分类。

6.根据权利要求1所述的具有信息安全保护模式的计算机存储系统,其特征在于:所述分类存储单元,包括:存储单元,用于将属性分类完成后将每个属性的数据存入对应的存储器中；检测将每个属性的数据存入对应的存储器的存储速率,并与预设存储速率进行比较,在确定存储速率小于预设存储速率时,发出报警提示；

所述存储单元,包括:解析模块和存储节点分配路径,存储节点分配路径由多个存储节点连接构成；所述解析模块与所述多个存储节点连接；

计算存储单元处理存储请求的数据量；

$$W_1 = \frac{\frac{\sum_{i=1}^N T_i}{N-1}}{\frac{T_0}{N} - \frac{\sum_{i=1}^N T_i}{N-1}}$$

其中, W_1 为存储单元处理存储请求的数据量; T_i 为第 i 个存储节点的存储速率; N 为节点的数量; T_0 为解析模块的分配速率;

计算存储单元在将对应属性的数据存入对应的存储器的过程中, 生成的数据量:

$$W_2 = \frac{\frac{1}{N} \left(Z - \frac{\sum_{i=1}^N F_i}{N-1} \right)}{\frac{Z}{N} - \frac{\sum_{i=1}^N F_i}{N-1}}$$

其中, W_2 为存储单元在将对应属性的数据存入对应的存储器的过程中, 生成的数据量; F_i 为第 i 个存储节点存储对应属性的数据后生成存储标签的速率; Z 为解析模块收回每个存储节点反馈的存储标签的速率;

根据存储单元处理存储请求的数据量及存储单元在将对应属性的数据存入对应的存储器的过程中生成的数据量, 计算将每个属性的数据存入对应的存储器的存储速率:

$$H = \frac{(W_1 + W_2) \times \sum_{i=1}^N T_i \times \sum_{i=1}^N F_i}{N^2 \sqrt{W_1^2 + W_2^2} \times \sqrt{\sum_{i=1}^N T_i^2 + \sum_{i=1}^N F_i^2}}$$

其中, H 为将每个属性的数据存入对应的存储器的存储速率。

具有信息安全保护模式的计算机存储系统

技术领域

[0001] 本发明涉及计算机存储技术领域,具体为具有信息安全保护模式的计算机存储系统。

背景技术

[0002] 计算机信息安全保护是确保组织内数据和系统免受潜在威胁的重要任务。

[0003] 公开号为CN115664857A的中国专利公开了一种网络信息安全的保护方法、系统、设备及存储介质,主要通过第一服务器向所述访问端推送虚拟数据信息,不仅能够避免非权限用户登录网络系统窃据信息,而且对于采用非法手段登录网络系统的用户进行识别,有效的防止机密信息被窃取,上述专利虽然解决了信息保护的问题,但是在实际操作中还存在以下问题:

[0004] 1.没有将存储器中的数据进行定期的周期性检测,从而使数据在进行多次读取后数据质量下降以及数据异常。

[0005] 2.没有将数据进行最大程度的数据加密,从而导致数据被破坏或者丢失。

[0006] 3.有根据数据的类型进行存储,并且数据存储时没有根据数据的大小选择合适的存储器,从而导致数据存储完整性不佳。

发明内容

[0007] 本发明的目的在于提供具有信息安全保护模式的计算机存储系统,通过对数据进行去除重复数据、填充缺失值、数据标准化的处理可以最大程度的保证数据的完整性,对周期序列数据和非周期序列数据中的异常数据都进行提取,可以快速定位和识别这些异常数据,以便及时采取措施进行处理和纠正,根据加密可以将数据转换为密文,只有持有密钥的人才能将其还原成原始数据。即使数据在传输过程中被窃取,盗取方也无法读取到原始数据,从而保护了数据的安全性,可以解决现有技术中的问题。

[0008] 为实现上述目的,本发明提供如下技术方案:

[0009] 具有信息安全保护模式的计算机存储系统,包括:

[0010] 身份验证单元,用于:

[0011] 进行系统操作时将操作者的管理身份进行身份认证,身份认证成功后操作者进行信息录入或者信息检索:

[0012] 分类存储单元,用于:

[0013] 基于身份验证单元中获取的录入信息,将录入信息根据关键词进行属性分类,属性分类完成后将每个属性的数据存入对应的存储器中,并将每个存储器中的数据进行数据备份;

[0014] 存储加密单元,用于:

[0015] 基于分类存储单元中获取的存储器中的数据,将存储器中的数据进行数据加密,并将加密完成的数据标注为加密数据;

[0016] 数据读取单元,用于:

[0017] 基于存储加密单元中获取的加密数据,将加密数据进行周期性检测,根据检测结果将加密数据中的异常数据进行确认,并将异常数据进行单独调取并显示;

[0018] 网络安全监测单元,用于:

[0019] 进行数据操作时对网络IP进行实时监测,若监测出异常IP时进行端口关闭,当网络IP进行正常信息操作时,进行实时网络流量监控分析,若通过分析网络流量数据有异常行为时,进行端口关闭。

[0020] 优选的,所述身份验证单元,还用于:

[0021] 操作者进行身份认证,身份认证为密码认证和账号认证,密码认证和账号认证后系统进行密码和账号的核对;

[0022] 核对完成后对该操作者的职位进行识别,其中,不同的职位进行不同的系统操作;

[0023] 身份认证成功后对操作者的网络IP进行验证,其中,将该操作者的网络IP与数据库中认证完成的网络IP进行核对,当操作者的网络IP与数据库中认证完成的网络IP一致时,该操作者可以进行下一步的信息数据操作。

[0024] 优选的,所述分类存储单元,包括:

[0025] 信息数据分类模块,用于:

[0026] 将录入的信息数据进行获取,并提取信息数据的关键信息片段;

[0027] 根据关键信息片段的目标语义确认录入信息数据的信息类型;

[0028] 其中,目标语义为将关键信息片段中位置连续的词语以文本最小单元进行最大长度的拼接,拼接后的得到拼接语义片段;

[0029] 根据词性统计特征对拼接语义片段进行清洗,得到片段关键词,其中,词性统计特征从数据库中进行获取;

[0030] 根据片段关键词确认录入信息数据的具体文本属性,并将具体文本数据作为该录入信息的信息类型。

[0031] 优选的,所述分类存储单元,还包括:

[0032] 分类数据存储模块,用于:

[0033] 基于信息数据分类模块中获取的录入信息数据的信息类型,将每个类型的录入信息数据进行长度确认;

[0034] 将录入信息数据划分为若干个长度相等的段落;

[0035] 根据段落的数量确认该录入信息数据的总长度,并根据录入信息数据的总长度计算出该录入信息数据的大小;

[0036] 根据录入信息数据的大小,将录入信息数据存储至对应的存储器中,其中,将存储器的剩余容量进行实时监控,录入信息数据的存量小于存储器的剩余容量;

[0037] 将每个存储器中的信息数据进行数据备份,并且每个备份数据均进行唯一编码标号。

[0038] 优选的,所述存储加密单元,还用于:

[0039] 将每个存储器中的录入信息数据进行加密段划分,并将每个划分完成的加密端进行加密序列建立;

[0040] 根据加密端的数量,将加密序列作为矩阵的第一列,建立第一加密矩阵;

- [0041] 根据预设干扰因子对每一待加密段在第一加密矩阵对应行上的第一位置进行调节,调节完毕后对其进行第一加密;
- [0042] 基于干扰次数在预设密码集中获取对应的干扰密码,同时利用加密骰子掷出随机数;
- [0043] 在第一加密矩阵的每一行中标记与第一位置相隔随机数的第二位置;
- [0044] 将干扰密码依次输入到第二位置进行第二加密,建立第二加密矩阵。
- [0045] 优选的,所述存储加密单元,还用于;
- [0046] 当第二加密矩阵中的空白位置的第一数量大于非空白位置的第二数量时,确定当前加密程度不足;
- [0047] 依次剔除第二矩阵的最外围位置,生成若干个子加密矩阵;
- [0048] 分别计算每一子加密矩阵的加密秩,基于所有的加密秩生成秩密码;
- [0049] 在第二加密矩阵上标记剩余空白位置,并将标记后的第加密矩阵与同型矩阵进行叠加,获取标记位置产生的叠加数;
- [0050] 将叠加数输入到剩余空白位置上,获取全加密矩阵;
- [0051] 提取第二全加密矩阵中每一位置对应的数据,获取加密数据。
- [0052] 优选的,所述数据读取单元,包括:
- [0053] 数据处理检测模块,用于:
- [0054] 将存储器中的加密数据进行获取,将获取的加密数据进行数据处理;
- [0055] 其中,数据处理包括去除重复数据、填充缺失值、数据标准化;
- [0056] 数据处理完成后进行周期性检测,根据周期性检测结果判断加密数据是否为周期序列数据。
- [0057] 优选的,所述数据读取单元,还包括:
- [0058] 数据分析模块,用于:
- [0059] 当数据进行周期性检测后为周期序列数据时,对周期序列数据按照周期进行分割得到多组相同的第一子序列,判断第一子序列中的全部数值是否大于预设数值,若是,提取出第一子序列中大于预设数值的第一异常数值,并将第一异常数值作为第一异常数据;
- [0060] 当数据进行周期性检测后为非周期序列数据时,通过一维聚类方法对周期序列数据进行聚类运算,得到多个分割点,并利用多个分割点对周期序列数据进行划分,得到多组不同的第二子序列,从第二子序列中获取存在大于预设数值的第三子序列,并将第三子序列作为第二异常数据;
- [0061] 将第一异常数据和第二异常数据进行数据整合,整合完成后将整合数据发送至显示终端进行数据对应显示,其中,数据对应为将整合数据存在的存储器以及录入网络IP进行对应。
- [0062] 优选的,所述分类存储单元,包括:分类单元,用于基于身份验证单元中获取的录入信息,将录入信息根据关键词进行属性分类;
- [0063] 所述分类单元,包括:
- [0064] 第一确定模块,用于:
- [0065] 将所述录入信息输入预先训练好的BERT模型中,输出录入信息包括的各个文本数据块的语义表示,根据语义表示结果确定文本数据块包括的关键词;

[0066] 计算每个文本数据块包括的关键词与预设数据库中各个预设关键词的匹配度：

$$[0067] \quad P_i = \frac{\sqrt{\left(\sum_{x=1}^N w_{i\partial_x}\right) \times \left(\sum_{y=1}^M \overline{w_{\varepsilon_y}}\right)}}{\sqrt{\left(\sum_{x=1}^N w_{i\partial_x}\right) \times \left(\sum_{y=1}^M \overline{w_{\varepsilon_y}}\right) + \sum_{x=1}^N \sum_{y=1}^M w_{i\partial_x} \times \overline{w_{\varepsilon_y}}}}$$

[0068] 其中, P_i 为每个文本数据块包括的关键词与预设数据库中第 i 个预设关键词的匹配度; ∂_x 为第 i 个预设关键词中第 x 个语义表示; ε_y 为文本数据块包括的关键词中第 y 个语义表示; $w_{i\partial_x}$ 为语义表示 ∂_x 在第 i 个预设关键词中出现的次数; $\overline{w_{\varepsilon_y}}$ 为语义表示 ε_y 在文本数据块包括的关键词中出现的次数; N 为第 i 个预设关键词中语义表示的数量; M 为文本数据块包括的关键词中语义表示的数量;

[0069] 根据最大匹配度对应的预设关键词确定每个文本数据库的数据类型, 并进行聚类, 得到若干个分类集合;

[0070] 第二确定模块, 用于:

[0071] 分别确定每个分类集合中的聚类中心, 确定各个聚类中心之间的距离;

[0072] 在确定各个聚类中心之间的距离均大于预设距离阈值时, 将若干个分类集合作为最终的属性分类结果; 反之, 重新将录入信息根据关键词进行属性分类。

[0073] 优选的, 所述分类存储单元, 包括: 存储单元, 用于将属性分类完成后将每个属性的数据存入对应的存储器中; 检测将每个属性的数据存入对应的存储器的存储速率, 并与预设存储速率进行比较, 在确定存储速率小于预设存储速率时, 发出报警提示;

[0074] 所述存储单元, 包括: 解析模块和存储节点分配路径, 存储节点分配路径由多个存储节点连接构成; 所述解析模块与所述多个存储节点连接;

[0075] 计算存储单元处理存储请求的数据量:

$$[0076] \quad W_1 = \frac{\sum_{i=1}^N T_i}{N-1} \cdot \frac{T_0}{N - \frac{\sum_{i=1}^N T_i}{N-1}}$$

[0077] 其中, W_1 为存储单元处理存储请求的数据量; T_i 为第 i 个存储节点的存储速率; N 为节点的数量; T_0 为解析模块的分配速率;

[0078] 计算存储单元在将对应属性的数据存入对应的存储器的过程中, 生成的数据量:

$$[0079] \quad W_2 = \frac{\frac{1}{N} \left(Z - \frac{\sum_{i=1}^N F_i}{N-1} \right)}{\frac{Z}{N} - \frac{\sum_{i=1}^N F_i}{N-1}}$$

[0080] 其中, W_2 为存储单元在将对应属性的数据存入对应的存储器的过程中, 生成的数

据量; F_i 为第*i*个存储节点存储对应属性的数据后生成存储标签的速率; Z 为解析模块收回每个存储节点反馈的存储标签的速率;

[0081] 根据存储单元处理存储请求的数据量及存储单元在将对应属性的数据存入对应的存储器的过程中生成的数据量,计算将每个属性的数据存入对应的存储器的存储速率:

$$[0082] \quad H = \frac{(W_1 + W_2) \times \sum_{i=1}^N T_i \times \sum_{i=1}^N F_i}{N^2 \sqrt{W_1^2 + W_2^2} \times \sqrt{\sum_{i=1}^N T_i^2 + \sum_{i=1}^N F_i^2}}$$

[0083] 其中, H 为将每个属性的数据存入对应的存储器的存储速率。

[0084] 与现有技术相比,本发明的有益效果如下:

[0085] 1.本发明提供的具有信息安全保护模式的计算机存储系统,通过关键词分类,可以将数据按照不同的主题和领域进行划分,方便用户根据需求快速定位和检索相关数据,根据录入信息数据的大小选择对应的存储器,可以最大程度的保证数据存储的完整性和安全性,同时将存储器中的数据再次进行数据备份可以将数据以副本的形式保存在安全的地方,以防止数据丢失、不可用或遭受破坏。

[0086] 2.本发明提供的具有信息安全保护模式的计算机存储系统,根据加密可以将数据转换为密文,只有持有密钥的人才能将其还原成原始数据。即使数据在传输过程中被窃取,盗取方也无法读取到原始数据,从而保护了数据的安全性。

[0087] 3.本发明提供的具有信息安全保护模式的计算机存储系统,通过对数据进行去除重复数据、填充缺失值、数据标准化的处理可以最大程度的保证数据的完整性,对周期序列数据和非周期序列数据中的异常数据都进行提取,可以快速定位和识别这些异常数据,以便及时采取措施进行处理和纠正。

附图说明

[0088] 图1为本发明的整体模块示意图;

[0089] 图2为本发明的整体方法示意图。

具体实施方式

[0090] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0091] 为了解决现有技术中,当操作者在进行信息录入时,没有对操作者的身份进行详细的验证,从而导致信息数据非法录入的问题,请参阅图1和图2,本实施例提供以下技术方案:

[0092] 具有信息安全保护模式的计算机存储系统,包括:

[0093] 身份验证单元,用于:

[0094] 进行系统操作时将操作者的管理身份进行身份认证,身份认证成功后操作者进行信息录入或者信息检索:

[0095] 分类存储单元,用于:

[0096] 基于身份验证单元中获取的录入信息,将录入信息根据关键词进行属性分类,属性分类完成后将每个属性的数据存入对应的存储器中,并将每个存储器中的数据进行数据备份;

[0097] 存储加密单元,用于:

[0098] 基于分类存储单元中获取的存储器中的数据,将存储器中的数据进行数据加密,并将加密完成的数据标注为加密数据;

[0099] 数据读取单元,用于:

[0100] 基于存储加密单元中获取的加密数据,将加密数据进行周期性检测,根据检测结果将加密数据中的异常数据进行确认,并将异常数据进行单独调取并显示;

[0101] 网络安全监测单元,用于:

[0102] 进行数据操作时对网络IP进行实时监测,若监测出异常IP时进行端口关闭,当网络IP进行正常信息操作时,进行实时网络流量监控分析,若通过分析网络流量数据有异常行为时,进行端口关闭。

[0103] 具体的,通过身份验证单元对每个操作者进行身份管理,可以最大程度的保证数据在录入和检索时的安全性,通过分类存储单元可以最大程度的保证数据存储的完整性和安全性,同时将存储器中的数据再次进行数据备份可以将数据以副本的形式保存在安全的地方,通过存储加密单元可以防止数据被篡改或破坏。使用加密技术可以保证数据的完整性,避免数据在传输过程中被非法修改或损坏,通过数据读取单元对周期序列数据和非周期序列数据中的异常数据都进行提取,可以快速定位和识别这些异常数据,以便及时采取措施进行处理和纠正,通过网络安全监测单元对数据进行网络安全监测,对进出网络的流量进行控制和监管,防止潜在的威胁和攻击,防止攻击扩散,从而可以提高数据输入和输出时的安全性。

[0104] 所述身份验证单元,还用于:

[0105] 操作者进行身份认证,身份认证为密码认证和账号认证,密码认证和账号认证后系统进行密码和账号的核对;

[0106] 核对完成后对该操作者的职位进行识别,其中,不同的职位进行不同的系统操作;

[0107] 身份认证成功后对操作者的网络IP进行验证,其中,将该操作者的网络IP与数据库中认证完成的网络IP进行核对,当操作者的网络IP与数据库中认证完成的网络IP一致时,该操作者可以进行下一步的信息数据操作。

[0108] 具体的,操作者先通过账号和密码进行身份认证,信息录入者需要进行身份注册,并且对每个操作者的职位进行管理,不同等级的职位所拥有的操作权限也是不同的,并且操作者在进行信息录入时,对操作者的网络IP进行核对,网络IP与数据库中认证完成的网络IP地址一致时,该操作者的网络IP才为安全的地址,若与数据库中认证完成的网络IP地址不一致时,操作者的数据则不进行录入。

[0109] 为了解决现有技术中,对录入的数据进行存储时,没有根据数据的类型进行存储,并且数据存储时没有根据数据的大小选择合适的存储器,从而导致数据存储完整性不佳的问题,请参阅图1和图2,本实施例提供以下技术方案:

[0110] 所述分类存储单元,包括:

- [0111] 信息数据分类模块,用于:
- [0112] 将录入的信息数据进行获取,并提取信息数据的关键信息片段;
- [0113] 根据关键信息片段的目标语义确认录入信息数据的信息类型;
- [0114] 其中,目标语义为将关键信息片段中位置连续的词语以文本最小单元进行最大长度的拼接,拼接后的得到拼接语义片段;
- [0115] 根据词性统计特征对拼接语义片段进行清洗,得到片段关键词,其中,词性统计特征从数据库中进行获取;
- [0116] 根据片段关键词确认录入信息数据的具体文本属性,并将具体文本数据作为该录入信息的信息类型。
- [0117] 分类数据存储模块,用于:
- [0118] 基于信息数据分类模块中获取的录入信息数据的信息类型,将每个类型的录入信息数据进行长度确认;
- [0119] 将录入信息数据划分为若干个长度相等的段落;
- [0120] 根据段落的数量确认该录入信息数据的总长度,并根据录入信息数据的总长度计算出该录入信息数据的大小;
- [0121] 根据录入信息数据的大小,将录入信息数据存储至对应的存储器中,其中,将存储器的剩余容量进行实时监控,录入信息数据的存量小于存储器的剩余容量;
- [0122] 将每个存储器中的信息数据进行数据备份,并且每个备份数据均进行唯一编码标号。
- [0123] 具体的,先通过信息数据分类模块将录入的信息数据进行数据分类,其中,根据关键信息片段的目标语义确认录入信息数据的信息类型,通过关键词分类,可以将数据按照不同的主题和领域进行划分,方便用户根据需求快速定位和检索相关数据。同时,对数据进行分类和组织,可以提高数据的可读性和可理解性,再通过分类数据存储模块将分类完成的数据存储至存储器中,并进行数据备份,其中,将存储器的剩余容量进行实时监控,录入信息数据的存量小于存储器的剩余容量,根据录入信息数据的大小选择对应的存储器,可以最大程度的保证数据存储的完整性和安全性,同时将存储器中的数据再次进行数据备份可以将数据以副本的形式保存在安全的地方,以防止数据丢失、不可用或遭受破坏。当原始数据发生问题时,备份数据可以提供及时、可靠的数据恢复,保障业务的连续性和数据的安全性。
- [0124] 为了解决现有技术中,数据进行存储后,没有将数据进行最大程度的数据加密,从而导致数据被破坏或者丢失的问题,请参阅图1和图2,本实施例提供以下技术方案:
- [0125] 所述存储加密单元,还用于:
- [0126] 将每个存储器中的录入信息数据进行加密段划分,并将每个划分完成的加密端进行加密序列建立;
- [0127] 根据加密端的数量,将加密序列作为矩阵的第一列,建立第一加密矩阵;
- [0128] 根据预设干扰因子对每一待加密段在第一加密矩阵对应行上的第一位置进行调节,调节完毕后对其进行第一加密;
- [0129] 基于干扰次数在预设密码集中获取对应的干扰密码,同时利用加密骰子掷出随机数;

- [0130] 在第一加密矩阵的每一行中标记与第一位置相隔随机数的第二位置;
- [0131] 将干扰密码依次输入到第二位置进行第二加密,建立第二加密矩阵。
- [0132] 当第二加密矩阵中的空白位置的第一数量大于非空白位置的第二数量时,确定当前加密程度不足;
- [0133] 依次剔除第二矩阵的最外围位置,生成若干个子加密矩阵;
- [0134] 分别计算每一子加密矩阵的加密秩,基于所有的加密秩生成秩密码;
- [0135] 在第二加密矩阵上标记剩余空白位置,并将标记后的第加密矩阵与同型矩阵进行叠加,获取标记位置产生的叠加数;
- [0136] 将叠加数输入到剩余空白位置上,获取全加密矩阵;
- [0137] 提取第二全加密矩阵中每一位置对应的数据,获取加密数据。
- [0138] 具体的,根据预设干扰因子对每一待加密段在第一加密矩阵对应行上的第一位置进行调节,调节完毕后对其进行第一加密,同时,根据干扰密码依次输入到第二位置进行第二加密,建立第二加密矩阵,根据加密可以将数据转换为密文,只有持有密钥的人才能将其还原成原始数据。即使数据在传输过程中被窃取,盗取方也无法读取到原始数据,从而保护了数据的安全性,同时数据加密技术可以防止数据被篡改或破坏。使用加密技术可以保证数据的完整性,避免数据在传输过程中被非法修改或损坏。
- [0139] 为了解决现有技术中,没有将存储器中的数据进行定期的周期性检测,从而使数据在进行多次读取后数据质量下降以及数据异常的问题,请参阅图1和图2,本实施例提供以下技术方案:
- [0140] 所述数据读取单元,包括:
- [0141] 数据处理检测模块,用于:
- [0142] 将存储器中的加密数据进行获取,将获取的加密数据进行数据处理;
- [0143] 其中,数据处理包括去除重复数据、填充缺失值、数据标准化;
- [0144] 数据处理完成后进行周期性检测,根据周期性检测结果判断加密数据是否为周期序列数据。
- [0145] 数据分析模块,用于:
- [0146] 当数据进行周期性检测后为周期序列数据时,对周期序列数据按照周期进行分割得到多组相同的第一子序列,判断第一子序列中的全部数值是否大于预设数值,若是,提取出第一子序列中大于预设数值的第一异常数值,并将第一异常数值作为第一异常数据;
- [0147] 当数据进行周期性检测后为非周期序列数据时,通过一维聚类方法对周期序列数据进行聚类运算,得到多个分割点,并利用多个分割点对周期序列数据进行划分,得到多组不同的第二子序列,从第二子序列中获取存在大于预设数值的第三子序列,并将第三子序列作为第二异常数据;
- [0148] 将第一异常数据和第二异常数据进行数据整合,整合完成后将整合数据发送至显示终端进行数据对应显示,其中,数据对应为将整合数据存在的存储器以及录入网络IP进行对应。
- [0149] 具体的,先通过数据处理检测模块对数据进行去除重复数据、填充缺失值、数据标准化的处理,通过对数据进行去除重复数据、填充缺失值、数据标准化的处理可以最大程度的保证数据的完整性,当保证数据的完整性后再对数据进行周期性检测,并且对周期序列

数据和非周期序列数据中的异常数据都进行提取,可以快速定位和识别这些异常数据,以便及时采取措施进行处理和纠正。这有助于避免数据质量下降和错误决策,并且数据周期性检测可以确保数据的准确性和完整性,从而提高数据质量。通过定期对数据进行校验和审核,可以发现并修复潜在的数据质量问题。

[0150] 优选的,所述分类存储单元,包括:分类单元,用于基于身份验证单元中获取的录入信息,将录入信息根据关键词进行属性分类;

[0151] 所述分类单元,包括:

[0152] 第一确定模块,用于:

[0153] 将所述录入信息输入预先训练好的BERT模型中,输出录入信息包括的各个文本数据块的语义表示,根据语义表示结果确定文本数据块包括的关键词;

[0154] 计算每个文本数据块包括的关键词与预设数据库中各个预设关键词的匹配度:

$$[0155] \quad P_i = \frac{\sqrt{\left(\sum_{x=1}^N w_{i\partial_x}\right) \times \left(\sum_{y=1}^M \overline{w_{\varepsilon_y}}\right)}}{\sqrt{\left(\sum_{x=1}^N w_{i\partial_x}\right) \times \left(\sum_{y=1}^M \overline{w_{\varepsilon_y}}\right) + \sum_{x=1}^N \sum_{y=1}^M w_{i\partial_x} \times \overline{w_{\varepsilon_y}}}}$$

[0156] 其中, P_i 为每个文本数据块包括的关键词与预设数据库中第*i*个预设关键词的匹配度; ∂_x 为第*i*个预设关键词中第*x*个语义表示; ε_y 为文本数据块包括的关键词中第*y*个语义表示; $w_{i\partial_x}$ 为语义表示 ∂_x 在第*i*个预设关键词中出现的次数; $\overline{w_{\varepsilon_y}}$ 为语义表示 ε_y 在文本数据块包括的关键词中出现的次数; N 为第*i*个预设关键词中语义表示的数量; M 为文本数据块包括的关键词中语义表示的数量;

[0157] 根据最大匹配度对应的预设关键词确定每个文本数据库的数据类型,并进行聚类,得到若干个分类集合;

[0158] 第二确定模块,用于:

[0159] 分别确定每个分类集合中的聚类中心,确定各个聚类中心之间的距离;

[0160] 在确定各个聚类中心之间的距离均大于预设距离阈值时,将若干个分类集合作为最终的属性分类结果;反之,重新将录入信息根据关键词进行属性分类。

[0161] 上述技术方案的工作原理:该实施例中,BERT模型为(Bidirectional Encoder Representations from Transformers):基于Transformer模型的语言模型,可以获得单词和文本的上下文相关表示。通过经过预训练的BERT模型,可以获取文本数据块的语义表示。

[0162] 该实施例中,根据语义表示结果确定文本数据块包括的关键词,包括:根据语义表示结果,利用自然语言处理的技术和算法,如TF-IDF(词频-逆文档频率)或基于机器学习的关键词提取模型,对文本数据块进行处理,提取出最具代表性和重要性的关键词。

[0163] 该实施例中,预设数据库中各个预设关键词为表示各个属性类型的关键词。

[0164] 该实施例中,计算每个文本数据块包括的关键词与预设数据库中各个预设关键词的匹配度,根据最大匹配度对应的预设关键词确定每个文本数据库的数据类型,将若干个文本数据块进行聚类,得到若干个分类集合。

[0165] 该实施例中,确定每个分类集合中的聚类中心,包括:对于每个分配到同一类别的

分类集合,计算分类集合中数据的平均值,作为聚类中心。

[0166] 该实施例中,预设距离阈值为表示不同类别的数据的差异距离。

[0167] 上述技术方案的有益效果:基于预先训练好的BERT模型,确定文本数据块包括的关键词,计算每个文本数据块包括的关键词与预设数据库中各个预设关键词的匹配度,进而便于准确实现根据关键词的聚类,得到若干个分类集合;计算各个聚类集合对应的聚类中心之间的距离,判断是否均大于预设距离阈值,在确定各个聚类中心之间的距离均大于预设距离阈值时,表示数据分类时合理的,提高数据分类的准确性。反之,重新将录入信息根据关键词进行属性分类。

[0168] 优选的,所述分类存储单元,包括:存储单元,用于将属性分类完成后将每个属性的数据存入对应的存储器中;检测将每个属性的数据存入对应的存储器的存储速率,并与预设存储速率进行比较,在确定存储速率小于预设存储速率时,发出报警提示;

[0169] 所述存储单元,包括:解析模块和存储节点分配路径,存储节点分配路径由多个存储节点连接构成;所述解析模块与所述多个存储节点连接;

[0170] 计算存储单元处理存储请求的数据量:

$$[0171] \quad W_1 = \frac{\frac{\sum_{i=1}^N T_i}{N-1}}{\frac{T_0}{N} - \frac{\sum_{i=1}^N T_i}{N-1}}$$

[0172] 其中, W_1 为存储单元处理存储请求的数据量; T_i 为第*i*个存储节点的存储速率; N 为节点的数量; T_0 为解析模块的分配速率;

[0173] 计算存储单元在将对应属性的数据存入对应的存储器的过程中,生成的数据量:

$$[0174] \quad W_2 = \frac{\frac{1}{N} \left(Z - \frac{\sum_{i=1}^N F_i}{N-1} \right)}{\frac{Z}{N} - \frac{\sum_{i=1}^N F_i}{N-1}}$$

[0175] 其中, W_2 为存储单元在将对应属性的数据存入对应的存储器的过程中,生成的数据量; F_i 为第*i*个存储节点存储对应属性的数据后生成存储标签的速率; Z 为解析模块收回每个存储节点反馈的存储标签的速率;

[0176] 根据存储单元处理存储请求的数据量及存储单元在将对应属性的数据存入对应的存储器的过程中生成的数据量,计算将每个属性的数据存入对应的存储器的存储速率:

$$[0177] \quad H = \frac{(W_1 + W_2) \times \sum_{i=1}^N T_i \times \sum_{i=1}^N F_i}{N^2 \sqrt{W_1^2 + W_2^2} \times \sqrt{\sum_{i=1}^N T_i^2 + \sum_{i=1}^N F_i^2}}$$

[0178] 其中, H 为将每个属性的数据存入对应的存储器的存储速率。

[0179] 上述技术方案的工作原理及有益效果:检测将每个属性的数据存入对应的存储器

的存储速率,并与预设存储速率进行比较,在确定存储速率小于预设存储速率时,发出报警提示,便于监测数据存储的存储速率,保证数据存储速率。在计算数据存储的存储速率的过程中,解析模块获取存储任务并进行解析,确定不同属性的数据对应的存储节点,基于存储节点分配路径进行数据存储分配,便于存储至相应的存储节点。在存储至相应的存储节点后,会生成相应的存储标签并反馈至解析模块,便于解析模块对各个存储节点的存储情况进行监控。在正向过程中,计算存储单元处理存储请求的数据量,在反馈过程中,计算存储单元在将对应属性的数据存入对应的存储器的过程中,生成的数据量。根据存储单元处理存储请求的数据量及存储单元在将对应属性的数据存入对应的存储器的过程中生成的数据量,准确计算出将每个属性的数据存入对应的存储器的存储速率,提高了判断存储速率与预设存储速率大小的准确性。

[0180] 需要说明的是,在本文中,诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来,而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、物品或者设备所固有的要素。

[0181] 尽管已经示出和描述了本发明的实施例,对于本领域的普通技术人员而言,可以理解在不脱离本发明的原理和精神的情况下可以对这些实施例进行多种变化、修改、替换和变型,本发明的范围由所附权利要求及其等同物限定。

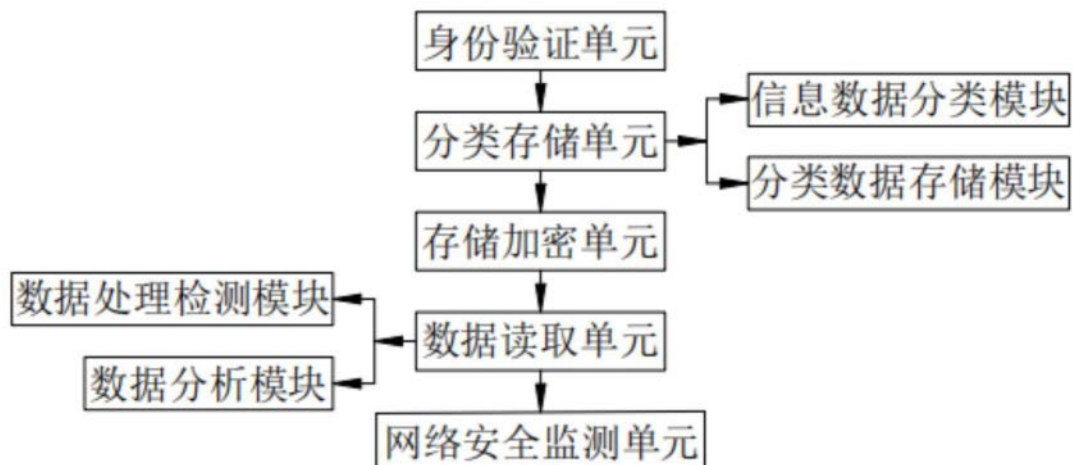


图1

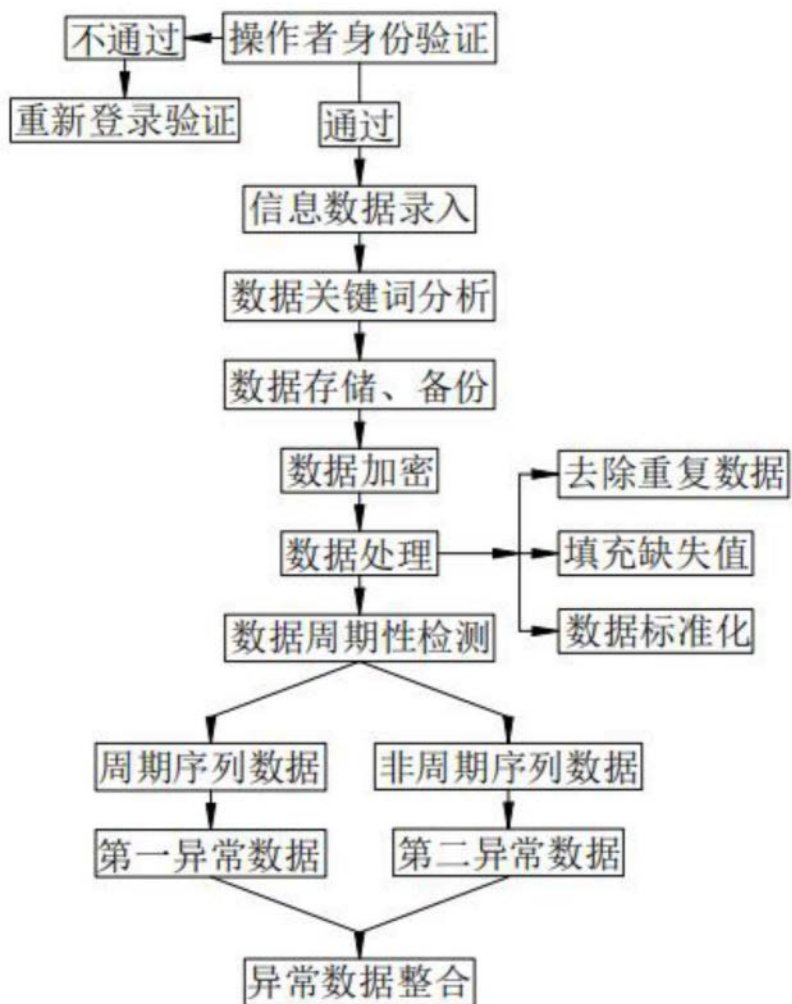


图2