(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 116796310 B
(45)授权公告日 2024.10.18

(21)申请号 202310703547.7

(22)申请日 2023.06.14

(65)同一申请的已公布的文献号
申请公布号 CN 116796310 A

(43)申请公布日 2023.09.22

(73)专利权人 上海瑞技计算机科技有限公司
地址 200030 上海市徐汇区银都路388号16幢316室

(72)发明人 郭萍 赵启东 包林陇

(74)专利代理机构 深圳市兰锋盛世知识产权代理有限公司 44504
专利代理师 王学

(51)Int.Cl.
G06F 21/55(2013.01)
G06F 21/56(2013.01)
G06F 16/36(2019.01)
G06F 16/35(2019.01)
G06F 18/23213(2023.01)

(56)对比文件
CN 111988339 A,2020.11.24
CN 114218568 A,2022.03.22

审查员 明媚

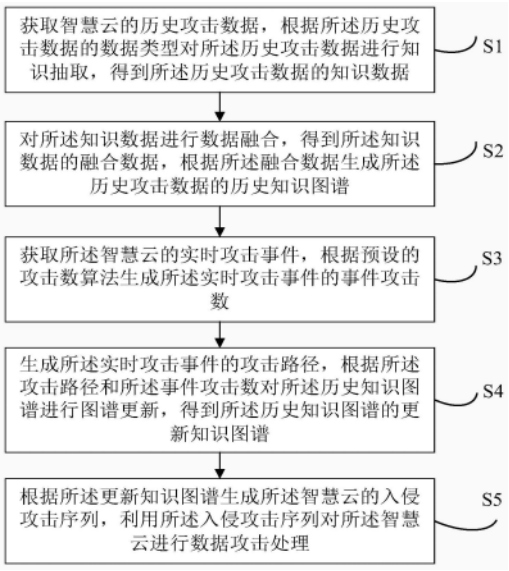
权利要求书3页 说明书11页 附图3页

(54)发明名称

应用于智慧云的数据攻击处理方法及系统

(57)摘要

本发明涉及人工智能技术领域,揭露了一种应用于智慧云的数据攻击处理方法及系统,包括:对智慧云的历史攻击数据进行知识抽取,得到历史攻击数据的知识数据;对知识数据进行数据融合,得到知识数据的融合数据,根据融合数据生成历史攻击数据的历史知识图谱;获取智慧云的实时攻击事件,根据预设的攻击数算法生成实时攻击事件的事件攻击数;生成实时攻击事件的攻击路径,根据攻击路径和事件攻击数对历史知识图谱进行图谱更新,得到历史知识图谱的更新知识图谱;根据更新知识图谱生成智慧云的入侵攻击序列,利用入侵攻击序列对智慧云进行数据攻击处理。本发明可以提高应用于智慧云的数据攻击处理的准确性。



CN 116796310 B

1.一种应用于智慧云的数据攻击处理方法,其特征在于,所述方法包括:

获取智慧云的历史攻击数据,根据所述历史攻击数据的数据类型对所述历史攻击数据进行知识抽取,得到所述历史攻击数据的知识数据;

对所述知识数据进行数据融合,得到所述知识数据的融合数据,根据所述融合数据生成所述历史攻击数据的历史知识图谱;

获取所述智慧云的实时攻击事件,根据预设的攻击数算法生成所述实时攻击事件的事件攻击数,其中,所述预设的攻击数算法为:

$$\begin{cases} \min\{J(U,V)\} = \sum_{k=1}^{n}\left\{\min\left[\sum_{i=1}^{c}(u_{ik})^{m}\parallel x_k - v_i \parallel^{2}\right]\right\} \\ \sum_{i=1}^{c} u_{ik} = 1 \end{cases}$$

其中,J(U,V)是各类中所述实时攻击事件的事件特征到聚类中心的距离平方和,min(*)是最小值函数,n是所述实时攻击事件的事件特征的特征总数,c是所述事件特征的指定分类数,u_{ik}是由所述时间特征产生的模糊分类矩阵中的第i行第k列的矩阵元素,m是程度系数,x_k是第k个所述事件特征,v_i是第i个所述聚类中心,i是所述事件特征的特征标识,k是所述聚类中心的类别标识,U是由所述时间特征产生的模糊分类矩阵,V是所述聚类中心的集合;

生成所述实时攻击事件的攻击路径,根据所述攻击路径和所述事件攻击数对所述历史知识图谱进行图谱更新,得到所述历史知识图谱的更新知识图谱;

根据所述更新知识图谱生成所述智慧云的入侵攻击序列,利用所述入侵攻击序列对所述智慧云进行数据攻击处理,其中,根据所述更新知识图谱生成所述智慧云的入侵攻击序列是因为所述更新知识图谱上包括所述智慧云的入侵攻击属性以及所述入侵攻击属性的属性值,其中入侵攻击序列是按照所述入侵攻击属性以及所述入侵攻击属性的属性值的大小进行排列;其中利用所述入侵攻击序列对所述智慧云进行数据攻击处理包括:根据所述入侵攻击序列中入侵攻击属性的属性值与预设的攻击阈值确定出需要进行处理的攻击,亦即,当所述属性值大于所述预设的攻击阈值时,根据所述属性值确定出对应的入侵攻击属性,以便准确的应对数据攻击。

2.如权利要求1所述的应用于智慧云的数据攻击处理方法,其特征在于,所述根据所述历史攻击数据的数据类型对所述历史攻击数据进行知识抽取,得到所述历史攻击数据的知识数据,包括:

根据所述历史攻击数据的数据来源确定所述历史攻击数据的数据格式,根据所述数据格式确定所述历史攻击数据的数据类型,其中,所述数据类型为结构化数据和非结构化数据;

对所述结构化数据进行三元组转换,得到所述结构化数据的三元组数据;

对所述非结构化数据进行信息抽取,得到所述非结构化数据的信息数据,汇集所述三元组数据和所述信息数据为所述历史攻击数据的知识数据。

3.如权利要求2所述的应用于智慧云的数据攻击处理方法,其特征在于,所述对所述结

构化数据进行三元组转换,得到所述结构化数据的三元组数据,包括:

对所述结构化数据进行数据映射,得到所诉结构化数据的映射数据;

根据预设的三元标签对所述映射数据进行数据选取,得到所述映射数据的目标数据;

确定所述目标数据的对应关系,根据所述对应关系和所述目标数据生成所述结构化数据的三元组数据。

4.如权利要求2所述的应用于智慧云的数据攻击处理方法,其特征在于,所述对所述非结构化数据进行信息抽取,得到所述非结构化数据的信息数据,包括:

对所述非结构化数据进行分词处理,得到所述非结构化数据的数据分词;

根据所述数据分词对所述非结构化数据进行本体抽取,得到所述非结构化数据的本体数据;

根据所述数据分词对所述非结构化数据进行实体抽取,得到所述非结构化数据的实体数据;

根据所述本体数据与所述实体数据生成所述非结构化数据的信息数据。

5.如权利要求1所述的应用于智慧云的数据攻击处理方法,其特征在于,所述对所述知识数据进行数据融合,得到所述知识数据的融合数据,包括:

对所述知识数据进行语义提取,得到所述知识数据的知识语义;

根据所述知识语义对所述知识数据的实体数据进行实体关联,得到所述实体数据的关联数据;

根据所述关联数据对所述知识数据进行数据融合,得到所述知识数据的融合数据。

6.如权利要求1所述的应用于智慧云的数据攻击处理方法,其特征在于,所述根据所述融合数据生成所述历史攻击数据的历史知识图谱,包括:

根据所述融合数据中的本体数据生成所述历史攻击数据的初始节点;

确定所述初始节点的节点标签,根据所述节点标签和所述融合数据对所述初始节点进行分支配置,得到所述初始节点的子节点;

根据所述初始节点和所述子节点生成所述历史攻击数据的历史知识图谱。

7.如权利要求1所述的应用于智慧云的数据攻击处理方法,其特征在于,所述根据预设的攻击数算法生成所述实时攻击事件的事件攻击数,包括:

对所述实时攻击事件进行特征提取,得到所述实时攻击事件的事件特征;

利用预设的攻击数算法对所述事件特征进行特征聚类,得到所述事件特征的聚类特征;

根据所述聚类特征生成所述实施攻击事件的事件攻击数。

8.如权利要求1所述的应用于智慧云的数据攻击处理方法,其特征在于,所述生成所述实时攻击事件的攻击路径,包括:

对所述实时攻击事件进行路径提取,得到所述实时攻击事件的事件路径;

对所述事件路径进行有效识别,得到所述事件路径的攻击路径。

9.如权利要求1至8中任一项所述的应用于智慧云的数据攻击处理方法,其特征在于,所述根据所述攻击路径和所述事件攻击数对所述历史知识图谱进行图谱更新,得到所述历史知识图谱的更新知识图谱,包括:

获取所述攻击路径的路径权重,利用预设的攻击加权算法、所述路径权重和所述事件

攻击数生成所述实时攻击事件的实时攻击权值,其中,所述预设的攻击加权算法为:

$$P=\sum_{j=1}^{T} w_j\delta_j$$

其中,P是所述实时攻击事件的实时攻击权值,w_j是第j个所述路径权重,δ_j是第j个所述路径权重所对应的所述事件攻击数,j是所述事件攻击数的标识,T是所述事件攻击数的总数;

利用所述实时攻击权值对所述历史知识图谱进行图谱更新,得到所述历史知识图谱的更新知识图谱。

10.一种应用于智慧云的数据攻击处理系统,其特征在于,用于执行如权利要求1-9中任一项所述的应用于智慧云的数据攻击处理方法,所述系统包括:

知识抽取模块,用于获取智慧云的历史攻击数据,根据所述历史攻击数据的数据类型对所述历史攻击数据进行知识抽取,得到所述历史攻击数据的知识数据;

数据融合模块,用于对所述知识数据进行数据融合,得到所述知识数据的融合数据,根据所述融合数据生成所述历史攻击数据的历史知识图谱;

攻击数生成模块,用于获取所述智慧云的实时攻击事件,根据预设的攻击数算法生成所述实时攻击事件的事件攻击数,其中,所述预设的攻击数算法为:

$$\begin{cases}\min\{J(U,V)\}=\sum_{k=1}^{n}\left\{\min\left[\sum_{i=1}^{c}(u_{ik})^m\parallel x_k-v_i\parallel^2\right]\right\}\\ \sum_{i=1}^{c}u_{ik}=1\end{cases}$$

其中,J(U,V)是各类中所述实时攻击事件的事件特征到聚类中心的距离平方和,min(*)是最小值函数,n是所述实时攻击事件的事件特征的特征总数,c是所述事件特征的指定分类数,u_{ik}是由所述时间特征产生的模糊分类矩阵中的第i行第k列的矩阵元素,m是程度系数,x_k是第k个所述事件特征,v_i是第i个所述聚类中心,i是所述事件特征的特征标识,k是所述聚类中心的类别标识,U是由所述时间特征产生的模糊分类矩阵,V是所述聚类中心的集合;

图谱更新模块,用于生成所述实时攻击事件的攻击路径,根据所述攻击路径和所述事件攻击数对所述历史知识图谱进行图谱更新,得到所述历史知识图谱的更新知识图谱;

攻击处理模块,用于根据所述更新知识图谱生成所述智慧云的入侵攻击序列,利用所述入侵攻击序列对所述智慧云进行数据攻击处理,其中,根据所述更新知识图谱生成所述智慧云的入侵攻击序列是因为所述更新知识图谱上包括所述智慧云的入侵攻击属性以及所述入侵攻击属性的属性值,其中入侵攻击序列是按照所述入侵攻击属性以及所述入侵攻击属性的属性值的大小进行排列;其中利用所述入侵攻击序列对所述智慧云进行数据攻击处理包括:根据所述入侵攻击序列中入侵攻击属性的属性值与预设的攻击阈值确定出需要进行处理的攻击,亦即,当所述属性值大于所述预设的攻击阈值时,根据所述属性值确定出对应的入侵攻击属性,以便准确的应对数据攻击。

应用于智慧云的数据攻击处理方法及系统

技术领域

[0001] 本发明涉及人工智能技术领域,尤其涉及一种应用于智慧云的数据攻击处理方法及系统。

背景技术

[0002] 随着数字化技术的不断进步,全球数字化发展已经进入加速阶段。在国家战略普遍支持和企业积极主动进行数字化转型的双驱动下,越来越多的企业将业务迁移到云上。智慧云足够便捷、方便、高性价比和弹性,但是一旦发生某智慧云泄漏了用户的数据隐私,或是数据在云端存储的过程中由于设备故障而导致大量丢失,亦或是数据在传输过程中被其他用户任意篡改,这种后果所造成的不良影响是难以估量,如何有效地实现数据攻击的防护处理是当下智慧云安全的工作重点。

[0003] 现如今,智慧云的安全防护存在一定的局限性,其攻击处理的精度难以得到保障,这样极容易产生错误的数据攻击处理,因此如何提升应用于智慧云的数据攻击处理时准确性,成为了亟待解决的问题。

发明内容

[0004] 本发明提供一种应用于智慧云的数据攻击处理方法及系统,其主要目的在于解决应用于智慧云的数据攻击处理时准确性较低的问题。

[0005] 为实现上述目的,本发明提供的一种应用于智慧云的数据攻击处理方法,包括:

[0006] 获取智慧云的历史攻击数据,根据所述历史攻击数据的数据类型对所述历史攻击数据进行知识抽取,得到所述历史攻击数据的知识数据;

[0007] 对所述知识数据进行数据融合,得到所述知识数据的融合数据,根据所述融合数据生成所述历史攻击数据的历史知识图谱;

[0008] 获取所述智慧云的实时攻击事件,根据预设的攻击数算法生成所述实时攻击事件的事件攻击数,其中,所述预设的攻击数算法为:

[0009] $$\begin{cases} \min\{J(U,V)\} = \sum_{k=1}^{n}\left\{\min\left[\sum_{i=1}^{c}(u_{ik})^{m}\parallel x_k - v_i \parallel^{2}\right]\right\} \\ \sum_{i=1}^{c} u_{ik} = 1 \end{cases}$$

[0010] 其中,J(U,V)是各类中所述实时攻击事件的事件特征到聚类中心的距离平方和,min(*)是最小值函数,n是所述实时攻击事件的事件特征的特征总数,c是所述事件特征的指定分类数,u_{ik}是由所述时间特征产生的模糊分类矩阵中的第i行第k列的矩阵元素m是程度系数,x_k是第k个所述事件特征,v_i是第i个所述聚类中心,i是所述事件特征的特征标识,k是所述聚类中心的类别标识,U是由所述时间特征产生的模糊分类矩阵,V是所述聚类中心的集合;

[0011] 生成所述实时攻击事件的攻击路径,根据所述攻击路径和所述事件攻击数对所述历史知识图谱进行图谱更新,得到所述历史知识图谱的更新知识图谱;

[0012] 根据所述更新知识图谱生成所述智慧云的入侵攻击序列,利用所述入侵攻击序列对所述智慧云进行数据攻击处理。

[0013] 可选地,所述根据所述历史攻击数据的数据类型对所述历史攻击数据进行知识抽取,得到所述历史攻击数据的知识数据,包括:

[0014] 根据所述历史攻击数据的数据来源确定所述历史攻击数据的数据格式,根据所述数据格式确定所述历史攻击数据的数据类型,其中,所述数据类型为结构化数据和非结构化数据;

[0015] 对所述结构化数据进行三元组转换,得到所述结构化数据的三元组数据;

[0016] 对所述非结构化数据进行信息抽取,得到所述非结构化数据的信息数据,汇集所述三元组数据和所述信息数据为所述历史攻击数据的知识数据。

[0017] 可选地,所述对所述结构化数据进行三元组转换,得到、所述结构化数据的三元组数据,包括:

[0018] 对所述结构化数据进行数据映射,得到所诉结构化数据的映射数据;

[0019] 根据预设的三元标签对所述映射数据进行数据选取,得到所述映射数据的目标数据;

[0020] 确定所述目标数据的对应关系,根据所述对应关系和所述目标数据生成所述结构化数据的三元组数据。

[0021] 可选地,所述对所述非结构化数据进行信息抽取,得到所述非结构化数据的信息数据,包括:

[0022] 对所述非结构化数据进行分词处理,得到所述非结构化数据的数据分词;

[0023] 根据所述数据分词对所述非结构化数据进行本体抽取,得到所述非结构化数据的本体数据;

[0024] 根据所述数据分词对所述非结构化数据进行实体抽取,得到所述非结构化数据的实体数据;

[0025] 根据所述本体数据与所述实体数据生成所述非结构化数据的信息数据。

[0026] 可选地,所述对所述知识数据进行数据融合,得到所述知识数据的融合数据,包括:

[0027] 对所述知识数据进行语义提取,得到所述知识数据的知识语义;

[0028] 根据所述知识语义对所述知识数据的实体数据进行实体关联,得到所述实体数据的关联数据;

[0029] 根据所述关联数据对所述知识数据进行数据融合,得到所述知识数据的融合数据。

[0030] 可选地,所述根据所述融合数据生成所述历史攻击数据的历史知识图谱,包括:

[0031] 根据所述融合数据中的本体数据生成所述历史攻击数据的初始节点;

[0032] 确定所述初始节点的节点标签,根据所述节点标签和所述融合数据对所述初始节点进行分支配置,得到所述初始节点的子节点;

[0033] 根据所述初始节点和所述子节点生成所述历史攻击数据的历史知识图谱。

[0034] 可选地,所述根据预设的攻击数算法生成所述实时攻击事件的事件攻击数,包括:

[0035] 对所述实时攻击事件进行特征提取,得到所述实时攻击事件的事件特征;

[0036] 利用预设的攻击数算法对所述事件特征进行特征聚类,得到所述事件特征的聚类特征;

[0037] 根据所述聚类特征生成所述实施攻击事件的事件攻击数。

[0038] 可选地,所述生成所述实时攻击事件的攻击路径,包括:

[0039] 对所述实时攻击事件进行路径提取,得到所述实时攻击事件的事件路径;

[0040] 对所述事件路径进行有效识别,得到所述事件路径的攻击路径。

[0041] 可选地,所述根据所述攻击路径和所述事件攻击数对所述历史知识图谱进行图谱更新,得到所述历史知识图谱的更新知识图谱,包括:

[0042] 获取所述攻击路径的路径权重,利用预设的攻击加权算法、所述路径权重和所述事件攻击数生成所述实时攻击事件的实时攻击权值,其中,所述预设的攻击加权算法为:

[0043] $$P=\sum_{j=1}^{T} w_j \delta_j$$

[0044] 其中,P是所述实时攻击事件的实时攻击权值,w_j是第j个所述路径权重,δ_j是第j个所述路径权重所对应的所述事件攻击数,j是所述事件攻击数的标识,T是所述事件攻击数的总数;

[0045] 利用所述实时攻击权值对所述历史知识图谱进行图谱更新,得到所述历史知识图谱的更新知识图谱。

[0046] 为了解决上述问题,本发明还提供一种应用于智慧云的数据攻击处理系统,所述系统包括:

[0047] 知识抽取模块,用于获取智慧云的历史攻击数据,根据所述历史攻击数据的数据类型对所述历史攻击数据进行知识抽取,得到所述历史攻击数据的知识数据;

[0048] 数据融合模块,用于对所述知识数据进行数据融合,得到所述知识数据的融合数据,根据所述融合数据生成所述历史攻击数据的历史知识图谱;

[0049] 攻击数生成模块,用于获取所述智慧云的实时攻击事件,根据预设的攻击数算法生成所述实时攻击事件的事件攻击数,其中,所述预设的攻击数算法为:

[0050] $$\begin{cases}\min\{J(U,V)\}=\sum_{k=1}^{n}\left\{\min\left[\sum_{i=1}^{c}(u_{ik})^m \| x_k-v_i \|^2\right]\right\}\\ \sum_{i=1}^{c} u_{ik}=1\end{cases}$$

[0051] 其中,J(U,V)是各类中所述实时攻击事件的事件特征到聚类中心的距离平方和,min(*)是最小值函数,n是所述实时攻击事件的事件特征的特征总数,c是所述事件特征的指定分类数,u_{ik}是由所述时间特征产生的模糊分类矩阵中的第i行第k列的矩阵元素m是程度系数,x_k是第k个所述事件特征,v_i是第i个所述聚类中心,i是所述事件特征的特征标识,k是所述聚类中心的类别标识,U是由所述时间特征产生的模糊分类矩阵,V是所述聚类中心

的集合;

[0052] 图谱更新模块,用于生成所述实时攻击事件的攻击路径,根据所述攻击路径和所述事件攻击数对所述历史知识图谱进行图谱更新,得到所述历史知识图谱的更新知识图谱;

[0053] 攻击处理模块,用于根据所述更新知识图谱生成所述智慧云的入侵攻击序列,利用所述入侵攻击序列对所述智慧云进行数据攻击处理。

[0054] 本发明实施例通过对智慧云的历史攻击数据进行知识抽取,得到所述历史攻击数据的知识数据,所述知识抽取是为了生成所述历史攻击数据的数据特征,对所述知识数据进行数据融合是为了根据数据融合生成的融合数据生成所述智慧云的历史攻击数据的历史知识图谱,所述历史知识图谱简单明了的显示了所述历史攻击数据的属性和属性值,利用实时攻击事件对所述历史知识图谱进行图谱更新,确保所述智慧云的知识图谱具有实时性,并且通过数据反馈形式可以使得所述更新知识图谱对所述智慧云的描述更加准确,根据所述更新知识图谱对所述智慧云进行数据攻击处理,是因为所述更新知识图谱显示了攻击程度和攻击路径,以便准确的应对攻击,因此本发明提出应用于智慧云的数据攻击处理方法及系统,可以解决应用于智慧云的数据攻击处理准确性较低的问题。

附图说明

[0055] 图1为本发明一实施例提供的应用于智慧云的数据攻击处理方法的流程示意图;

[0056] 图2为本发明一实施例提供的生成历史攻击数据的知识数据的流程示意图;

[0057] 图3为本发明一实施例提供的生成非结构化数据的信息数据的流程示意图;

[0058] 图4为本发明一实施例提供的应用于智慧云的数据攻击处理系统的功能模块图;

[0059] 本发明目的的实现、功能特点及优点将结合实施例,参照附图做进一步说明。

具体实施方式

[0060] 应当理解,此处所描述的具体实施例仅仅用以解释本发明,并不用于限定本发明。

[0061] 本申请实施例提供一种应用于智慧云的数据攻击处理方法。所述应用于智慧云的数据攻击处理方法的执行本体包括但不限于服务端、终端等能够被配置为执行本申请实施例提供的该方法的电子设备中的至少一种。换言之,所述应用于智慧云的数据攻击处理方法可以由安装在终端设备或服务端设备的软件或硬件来执行,所述软件可以是区块链平台。所述服务端包括但不限于:单台服务器、服务器集群、云端服务器或云端服务器集群等。所述服务器可以是独立的服务器,也可以是提供云服务、云数据库、云计算、云函数、云存储、网络服务、云通信、中间件服务、域名服务、安全服务、内容分发网络(Content Delivery Network,CDN)、以及大数据和人工智能平台等基础云计算服务的云服务器。

[0062] 参照图1所示,为本发明一实施例提供的应用于智慧云的数据攻击处理方法的流程示意图。在本实施例中,所述应用于智慧云的数据攻击处理方法包括:

[0063] S1、获取智慧云的历史攻击数据,根据所述历史攻击数据的数据类型对所述历史攻击数据进行知识抽取,得到所述历史攻击数据的知识数据。

[0064] 在本发明实施例中,所述智慧云可以有效管理企业的信息,企业的很多业务可以利用智慧云进行管理,例如:办公业务、销售业务、采购业务、库存业务、客户管理、项目管

理、人事管理、培训管理、成本管理、关键业绩指标等。

[0065] 详细地,所述获取智慧云的历史攻击数据可以是本领域习知的消息采集工具(例如,可以是Flume、Kafka等工具),所述Flume和Kafka工具仅仅是为了说明方案的可实施性而进行的示例性举例,并不限定本方案必须采用所述Flume和Kafka工具。

[0066] 详细地,Kafka是由Apache软件基金会开发的一个开源流处理平台,是一种高吞吐量的分布式发布订阅消息系统,它可以处理消费者在网站中的所有动作流数据。

[0067] 详细地,所述数据类型分为结构化数据和非结构化数据,所述非结构化数据分为半结构化数据和纯文本数据,其中,所述结构化数据一般是指可以使用关系型数据库表示和存储,可以用二维表来逻辑表达实现的数据,所述非结构化数据可以来自研究报告、学术期刊和公开博客等;所述结构化数据来自各大数据库。

[0068] 在本发明实施例中,如图2所示,所述根据所述历史攻击数据的数据类型对所述历史攻击数据进行知识抽取,得到所述历史攻击数据的知识数据,包括:

[0069] S21、根据所述历史攻击数据的数据来源确定所述历史攻击数据的数据格式,根据所述数据格式确定所述历史攻击数据的数据类型,其中,所述数据类型为结构化数据和非结构化数据;

[0070] S22、对所述结构化数据进行三元组转换,得到所述结构化数据的三元组数据;

[0071] S23、对所述非结构化数据进行信息抽取,得到所述非结构化数据的信息数据,汇集所述三元组数据和所述信息数据为所述历史攻击数据的知识数据。

[0072] 详细地,所述历史攻击数据的数据来源包括所述数据库、研究报告、学术期刊和公开博客等;所述数据格式包括:表格、图片文件、纯文件、二进制文件、音频格式以及视频格式等。

[0073] 详细地,所述非结构化数据是数据结构不规则或不完整,没有预定义的数据模型,不方便用数据库二维逻辑表来表现的数据,包括所有格式的办公文档、文本、图片,HTML、各类报表、图像和音频/视频信息等等。

[0074] 详细地,所述对所述结构化数据进行三元组转换,得到、所述结构化数据的三元组数据,包括:

[0075] 对所述结构化数据进行数据映射,得到所诉结构化数据的映射数据;

[0076] 根据预设的三元标签对所述映射数据进行数据选取,得到所述映射数据的目标数据;

[0077] 确定所述目标数据的对应关系,根据所述对应关系和所述目标数据生成所述结构化数据的三元组数据。

[0078] 详细地,所述根据预设的三元标签是指所述映射数据的属性、所述映射数据的属性值以及所述映射数据的资源,其中,所述三元组数据可以利用表格进行表征,例如:表的列作为所述映射数据的属性、表的行作为所述映射数据的资源、表的单元格值为所述映射数据字面量,可以用表作为所述结构化数据中本体中的类,所述资源又称为所述映射数据中的实体。

[0079] 进一步地,可以将所述三元组数据用节点和边表示,其中,所述节点表示所述映射数据中的实体和属性,边则表示了实体和实体之间的关系以及实体和属性的对应关系,其中,所述目标数据的对应关系是指实体和实体之间的关系以及实体和属性之间的关系。

[0080] 详细地,所述三元组数据用来表示实体与实体之间的关系,或者实体的某个属性的属性值是什么,从内容上看三元组的结构为“资源-属性-属性值”,实体由统一资源标识符表示,属性值可以是另一个实体的统一资源标识符,也可以是某种数据类型的值,也称为字面量。

[0081] 详细地,参图3所示,所述对所述非结构化数据进行信息抽取,得到所述非结构化数据的信息数据,包括:

[0082] S31、对所述非结构化数据进行分词处理,得到所述非结构化数据的数据分词;

[0083] S32、根据所述数据分词对所述非结构化数据进行本体抽取,得到所述非结构化数据的本体数据;

[0084] S33、根据所述数据分词对所述非结构化数据进行实体抽取,得到所述非结构化数据的实体数据;

[0085] S34、根据所述本体数据与所述实体数据生成所述非结构化数据的信息数据。

[0086] 详细地,所述对所述非结构化数据进行分词处理可以是本领域习知的分词工具(例如,可以是jieba、SnowNLP、PkuSeg、THULAC、HanLP等工具),所述jieba、SnowNLP、PkuSeg、THULAC、HanLP工具仅仅是为了说明方案的可实施性而进行的示例性举例,并不限定本方案必须采用所述jieba、SnowNLP、PkuSeg、THULAC、HanLP工具。

[0087] 详细地,所述根据所述数据分词对所述非结构化数据进行本体抽取是指确定所述非结构化数据中的的本体数据,其中,本体是概念的集合,是公认的概念框架,一般指不会改变如“人”、“事”、“物”、“地”、“组织”,在面对对象编程里面,可以将所述本体称为类,在数据管理里面,可以将所述本体称为元数据。

[0088] 详细地,所述根据所述数据分词对所述非结构化数据进行实体抽取是指确定出所述非结构化数据中的实体数据,其中,所述实体是本体、实例及关系的整合,比如“人”是本体框中的一个概念,概念中也规定了相关属性比如“性别”,小明是一个具体的人,叫做实例,所以小明也有性别,小明以及体现小明的本体概念“人”以及相关属性,叫做一个实体。

[0089] S2、对所述知识数据进行数据融合,得到所述知识数据的融合数据,根据所述融合数据生成所述历史攻击数据的历史知识图谱。

[0090] 在本发明实施例中,所述历史知识图谱是对所述历史攻击数据的关联关系以及数据权重的表示,其中,知识图谱是一种图谱组织形式,通过语义关联把各种实体关联起来,知识图谱把结构化、非结构化的数据通过数据抽取、融合在一起,有利于大规模数据的利用和迁移。

[0091] 详细地,所述历史攻击数据的历史知识图谱是根据提取出所述历史攻击数据的实体、属性和关系等信息,然后再经过知识融合步骤,获得消除歧义、关系更明朗的高质量数据。

[0092] 在本发明实施例中,所述对所述知识数据进行数据融合,得到所述知识数据的融合数据,包括:

[0093] 对所述知识数据进行语义提取,得到所述知识数据的知识语义;

[0094] 根据所述知识语义对所述知识数据的实体数据进行实体关联,得到所述实体数据的关联数据;

[0095] 根据所述关联数据对所述知识数据进行数据融合,得到所述知识数据的融合数

据。

[0096] 详细地,所述对所述知识数据进行语义提取可以利用bert模型。

[0097] 详细地,所述根据所述知识语义对所述知识数据的实体数据进行实体关联是指根据所述知识语义确定所述实体与本体的关系,以及所述实体与实体之间的关系。

[0098] 详细地,所述根据所述关联数据对所述知识数据进行数据融合可以利用贝叶斯估计法,所述贝叶斯估计法是根据观测空间的先验知识,提供一种计算后验概率的方法,实现观测空间中目标的识别,所述贝叶斯估计法易于理解,且计算量小;所述根据所述关联数据对所述知识数据进行数据融合也可以利用极大似然估计算法,其中,将所述融合数据取为使似然函数达到极值的估计值,所述极大似然估计算法的信息丢失少,适合对所述知识数据进行融合。

[0099] 在本发明实施例中,所述根据所述融合数据生成所述历史攻击数据的历史知识图谱,包括:

[0100] 根据所述融合数据中的本体数据生成所述历史攻击数据的初始节点;

[0101] 确定所述初始节点的节点标签,根据所述节点标签和所述融合数据对所述初始节点进行分支配置,得到所述初始节点的子节点;

[0102] 根据所述初始节点和所述子节点生成所述历史攻击数据的历史知识图谱。

[0103] 详细地,所述根据所述融合数据中的本体数据生成所述历史攻击数据的初始节点是指根据所述本体数据确定所述历史攻击数据的分类类别,所述本体数据所表征的本体是所述历史攻击数据的分类类别;所述分类类别表示所述历史知识图谱的本体,所述确定所述初始节点的节点标签是指为所述初始节点赋予一个标识性的标签,所述节点标签用来区分不同的初始节点,用来区分不同的本体。

[0104] 详细地,所述本体是对特定领域之中某套概念及其相互之间关系的形式化表达,所述历史知识图谱的本体可以理解为知识图谱的数据模式,通常也可以描述为一个语义网络,描述了所述历史知识图谱中有哪些类型的实体,每个类型的实体有哪些类型的属性以及各类实体之间有哪些类型的关系。

[0105] 详细地,所述根据所述节点标签和所述融合数据对所述初始节点进行分支配置是指根据融合数据确定与本体关联的实体的属性,其中,所述实体指具有可区别性且独立存在的某种事物,是所述历史知识图谱的最基本元素,在所述历史知识图谱中表现为语义网络中的顶点;所述属性指描述事物自身性质或者事物间联系的性质的信息,在所述历史知识图谱中表现为实体的键值对数据,所述键值对是由属性名和属性值组成。

[0106] 详细地,所述根据所述初始节点和所述子节点生成所述历史攻击数据的历史知识图谱是指根据所述初始节点和所述子节点确定所述历史攻击数据的实体、属性、关系等元素,利用实体、属性、关系等元素生成所述历史攻击数据的历史知识图谱,其中,所述关系是指事物之间普遍存在的联系,在所述历史知识图谱中就是描述实体之间联系的语义边。

[0107] S3、获取所述智慧云的实时攻击事件,根据预设的攻击数算法生成所述实时攻击事件的事件攻击数。

[0108] 在本发明实施例中,所述根据预设的攻击数算法生成所述实时攻击事件的事件攻击数,包括:

[0109] 对所述实时攻击事件进行特征提取,得到所述实时攻击事件的事件特征;

[0110] 利用预设的攻击数算法对所述事件特征进行特征聚类,得到所述事件特征的聚类特征;

[0111] 根据所述聚类特征生成所述实施攻击事件的事件攻击数。

[0112] 详细地,所述对所述实时攻击事件进行特征提取可以先对所述实时攻击事件先进行格式转化,再将格式转化后的数据进行数据分词,对分词后的数据进行数据选取,得到所述实时攻击事件的事件特征。

[0113] 详细地,所述预设的攻击数算法为:

[0114] $$\begin{cases} \min\{J(U,V)\} = \sum_{k=1}^{n}\left\{\min\left[\sum_{i=1}^{c}(u_{ik})^{m}\parallel x_k - v_i \parallel^{2}\right]\right\} \\ \sum_{i=1}^{c} u_{ik} = 1 \end{cases}$$

[0115] 其中,J(U,V)是各类中所述实时攻击事件的事件特征到聚类中心的距离平方和,min(*)是最小值函数,n是所述实时攻击事件的事件特征的特征总数,c是所述事件特征的指定分类数,u_{ik}是由所述时间特征产生的模糊分类矩阵中的第i行第k列的矩阵元素m是程度系数,x_k是第k个所述事件特征,v_i是第i个所述聚类中心,i是所述事件特征的特征标识,k是所述聚类中心的类别标识,U是由所述时间特征产生的模糊分类矩阵,V是所述聚类中心的集合。

[0116] 详细地,根据所述预设的攻击数算法确定在约束条件下所述实时攻击事件的事件特征到聚类中心的距离平方和的最小值,根据最小值进行所述事件特征的特征聚类,因为所述最小值表示所述事件特征与所述聚类中心的相似性。

[0117] 详细地,所述根据所述聚类特征生成所述实施攻击事件的事件攻击数是指根据所述聚类特征确定某一聚类中心下的事件特征的个数,确定某一聚类中心下的事件特征的个数为所述实施攻击事件生成的其中一类攻击的的事件攻击数。

[0118] S4、生成所述实时攻击事件的攻击路径,根据所述攻击路径和所述事件攻击数对所述历史知识图谱进行图谱更新,得到所述历史知识图谱的更新知识图谱。

[0119] 在本发明实施例中,所述生成所述实时攻击事件的攻击路径,包括:

[0120] 对所述实时攻击事件进行路径提取,得到所述实时攻击事件的事件路径;

[0121] 对所述事件路径进行有效识别,得到所述事件路径的攻击路径。

[0122] 详细地,所述对所述实时攻击事件进行路径提取可以根据页面插件确定所述实施攻击事件的事件日志,对所述事件日志进行分词处理,得到所述事件日志的日志分词,对所述日志分词进行分词选取,得到所述实时攻击事件的事件路径。

[0123] 详细地,所述对所述事件路径进行有效识别是指针对同一种事件,所述事件的事件路径不同,那么事件的安全性是不同的,例如,当所述登录事件是从常用地址发出的,那么,可以初步确定所述登录事件是正常的,当登录是异地发生的,需要进一步确定登录事件是否正常,亦即,所述有效识别就是异常判断。

[0124] 在本发明实施例中,所述根据所述攻击路径和所述事件攻击数对所述历史知识图谱进行图谱更新,得到所述历史知识图谱的更新知识图谱,包括:

[0125] 获取所述攻击路径的路径权重,利用预设的攻击加权算法、所述路径权重和所述

事件攻击数生成所述实时攻击事件的实时攻击权值,其中,所述预设的攻击加权算法为:

[0126] $$P=\sum_{j=1}^{T} w_j\delta_j$$

[0127] 其中,P是所述实时攻击事件的实时攻击权值,w_j是第j个所述路径权重,δ_j是第j个所述路径权重所对应的所述事件攻击数,j是所述事件攻击数的标识,T是所述事件攻击数的总数;

[0128] 利用所述实时攻击权值对所述历史知识图谱进行图谱更新,得到所述历史知识图谱的更新知识图谱。

[0129] 详细地,所述攻击路径的路径权重是指根据所述攻击路径的路径风险程度确定的;所述利用所述实时攻击权值对所述历史知识图谱进行图谱更新是指根据所述实时攻击权值对所述历史知识图谱中实体的属性、所述实体的属性值、以及所述属性的连接关系进行更新。

[0130] S5、根据所述更新知识图谱生成所述智慧云的入侵攻击序列,利用所述入侵攻击序列对所述智慧云进行数据攻击处理。

[0131] 在本发明实施例中,所述根据所述更新知识图谱生成所述智慧云的入侵攻击序列是因为所述更新知识图谱上显示了所述智慧云的入侵攻击属性以及所述入侵攻击属性的属性值,所述入侵攻击序列是按照所述入侵攻击属性以及所述入侵攻击属性的属性值的大小进行排列的。

[0132] 在本发明实施例中,所述利用所述入侵攻击序列对所述智慧云进行数据攻击处理是指根据所述入侵攻击序列中入侵攻击属性的属性值与预设的攻击阈值确定出需要进行处理的攻击,亦即,当所述属性值大于所述预设的攻击阈值时,根据所述属性值确定出对应的入侵攻击属性,根据所述入侵攻击属性对所述智慧云进行相应的数据攻击处理。

[0133] 本发明实施例通过对智慧云的历史攻击数据进行知识抽取,得到所述历史攻击数据的知识数据,所述知识抽取是为了生成所述历史攻击数据的数据特征,对所述知识数据进行数据融合是为了根据数据融合生成的融合数据生成所述智慧云的历史攻击数据的历史知识图谱,所述历史知识图谱简单明了的显示了所述历史攻击数据的属性和属性值,利用实时攻击事件对所述历史知识图谱进行图谱更新,确保所述智慧云的知识图谱具有实时性,并且通过数据反馈形式可以使得所述更新知识图谱对所述智慧云的描述更加准确,根据所述更新知识图谱对所述智慧云进行数据攻击处理,是因为所述更新知识图谱显示了攻击程度和攻击路径,以便准确的应对攻击,因此本发明提出应用于智慧云的数据攻击处理方法,可以解决应用于智慧云的数据攻击处理准确性较低的问题。

[0134] 如图4所示,是本发明一实施例提供的应用于智慧云的数据攻击处理系统的功能模块图。

[0135] 本发明所述应用于智慧云的数据攻击处理系统100可以安装于电子设备中。根据实现的功能,所述应用于智慧云的数据攻击处理系统100可以包括知识抽取模块101、数据融合模块102、攻击数生成模块103、图谱更新模块104及攻击处理模块105。本发明所述模块也可以称之为单元,是指一种能够被电子设备处理器所执行,并且能够完成固定功能的一系列计算机程序段,其存储在电子设备的存储器中。

[0136] 在本实施例中,关于各模块/单元的功能如下:

[0137] 所述知识抽取模块101,用于获取智慧云的历史攻击数据,根据所述历史攻击数据的数据类型对所述历史攻击数据进行知识抽取,得到所述历史攻击数据的知识数据;

[0138] 所述数据融合模块102,用于对所述知识数据进行数据融合,得到所述知识数据的融合数据,根据所述融合数据生成所述历史攻击数据的历史知识图谱;

[0139] 所述攻击数生成模块103,用于获取所述智慧云的实时攻击事件,根据预设的攻击数算法生成所述实时攻击事件的事件攻击数,其中,所述预设的攻击数算法为:

[0140] $$\begin{cases} \min\{J(U,V)\} = \sum_{k=1}^{n}\left\{\min\left[\sum_{i=1}^{c}(u_{ik})^{m}\parallel x_k - v_i \parallel^{2}\right]\right\} \\ \sum_{i=1}^{c} u_{ik} = 1 \end{cases}$$

[0141] 其中,J(U,V)是各类中所述实时攻击事件的事件特征到聚类中心的距离平方和,min(*)是最小值函数,n是所述实时攻击事件的事件特征的特征总数,c是所述事件特征的指定分类数,u_{ik}是由所述时间特征产生的模糊分类矩阵中的第i行第k列的矩阵元素m是程度系数,x_k是第k个所述事件特征,v_i是第i个所述聚类中心,i是所述事件特征的特征标识,k是所述聚类中心的类别标识,U是由所述时间特征产生的模糊分类矩阵,V是所述聚类中心的集合;

[0142] 所述图谱更新模块104,用于生成所述实时攻击事件的攻击路径,根据所述攻击路径和所述事件攻击数对所述历史知识图谱进行图谱更新,得到所述历史知识图谱的更新知识图谱;

[0143] 所述攻击处理模块105,用于根据所述更新知识图谱生成所述智慧云的入侵攻击序列,利用所述入侵攻击序列对所述智慧云进行数据攻击处理。

[0144] 在本发明所提供的几个实施例中,应该理解到,所揭露的方法和系统,可以通过其它的方式实现。例如,以上所描述的系统实施例仅仅是示意性的,例如,所述模块的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式。

[0145] 所述作为分离部件说明的模块可以是或者也可以不是物理上分开的,作为模块显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。

[0146] 另外,在本发明各个实施例中的各功能模块可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现,也可以采用硬件加软件功能模块的形式实现。

[0147] 对于本领域技术人员而言,显然本发明不限于上述示范性实施例的细节,而且在不背离本发明的精神或基本特征的情况下,能够以其他的具体形式实现本发明。

[0148] 因此,无论从哪一点来看,均应将实施例看作是示范性的,而且是非限制性的,本发明的范围由所附权利要求而不是上述说明限定,因此旨在将落在权利要求的等同要件的含义和范围内的所有变化涵括在本发明内。不应将权利要求中的任何附关联图标记视为限制所涉及的权利要求。

[0149] 本发明所指区块链是分布式数据存储、点对点传输、共识机制、加密算法等计算机技术的新型应用模式。区块链(Blockchain),本质上是一个去中心化的数据库,是一串使用密码学方法相关联产生的数据块,每一个数据块中包含了一批次网络交易的信息,用于验证其信息的有效性(防伪)和生成下一个区块。区块链可以包括区块链底层平台、平台产品服务层以及应用服务层等。

[0150] 本申请实施例可以基于人工智能技术对相关的数据进行获取和处理。其中,人工智能(Artificial Intelligence,AI)是利用数字计算机或者数字计算机控制的机器模拟、延伸和扩展人的智能,感知环境、获取知识并使用知识获得最佳结果的理论、方法、技术及应用系统。

[0151] 此外,显然“包括”一词不排除其他单元或步骤,单数不排除复数。系统权利要求中陈述的多个单元或系统也可以由一个单元或系统通过软件或者硬件来实现。第一、第二等词语用来表示名称,而并不表示任何特定的顺序。

[0152] 最后应说明的是,以上实施例仅用以说明本发明的技术方案而非限制,尽管参照较佳实施例对本发明进行了详细说明,本领域的普通技术人员应当理解,可以对本发明的技术方案进行修改或等同替换,而不脱离本发明技术方案的精神和范围。

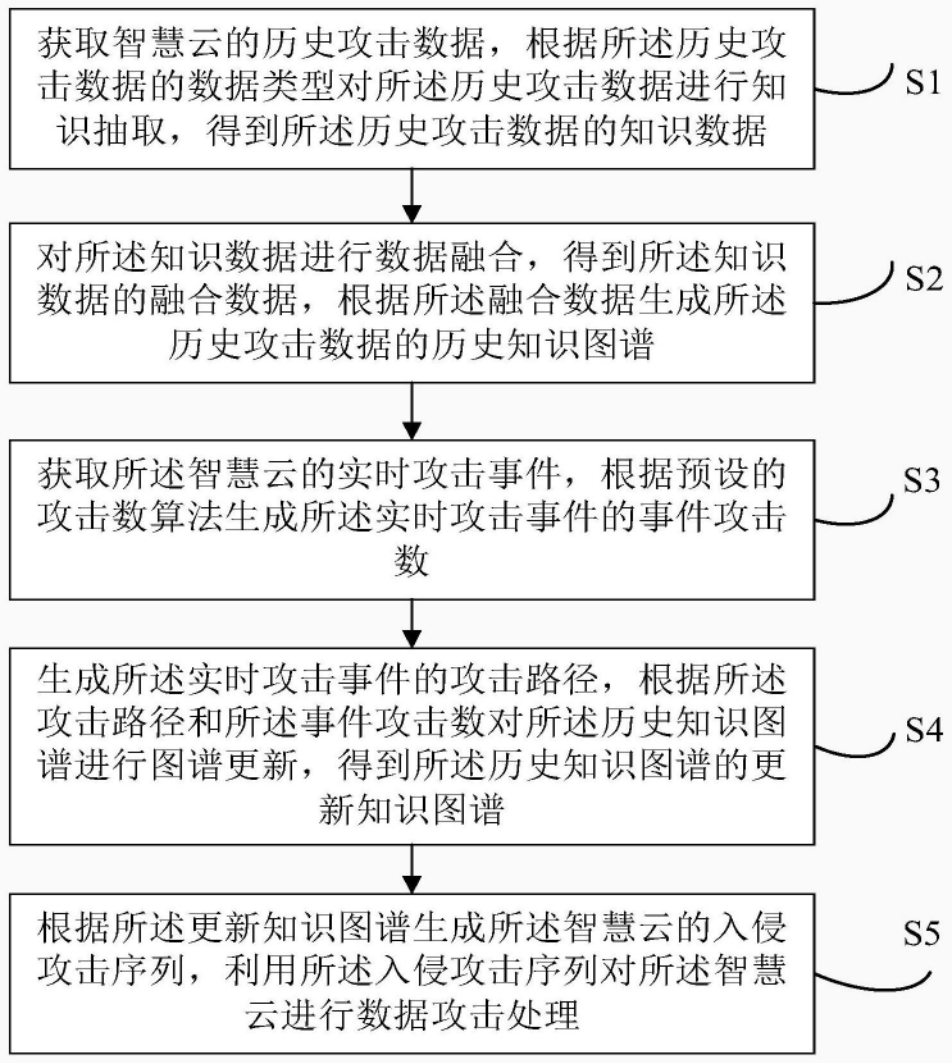

获取智慧云的历史攻击数据，根据所述历史攻击数据的数据类型对所述历史攻击数据进行知识抽取，得到所述历史攻击数据的知识数据
S1
对所述知识数据进行数据融合，得到所述知识数据的融合数据，根据所述融合数据生成所述历史攻击数据的历史知识图谱
S2
获取所述智慧云的实时攻击事件，根据预设的攻击数算法生成所述实时攻击事件的事件攻击数
S3
生成所述实时攻击事件的攻击路径，根据所述攻击路径和所述事件攻击数对所述历史知识图谱进行图谱更新，得到所述历史知识图谱的更新知识图谱
S4
根据所述更新知识图谱生成所述智慧云的入侵攻击序列，利用所述入侵攻击序列对所述智慧云进行数据攻击处理
S5

图1

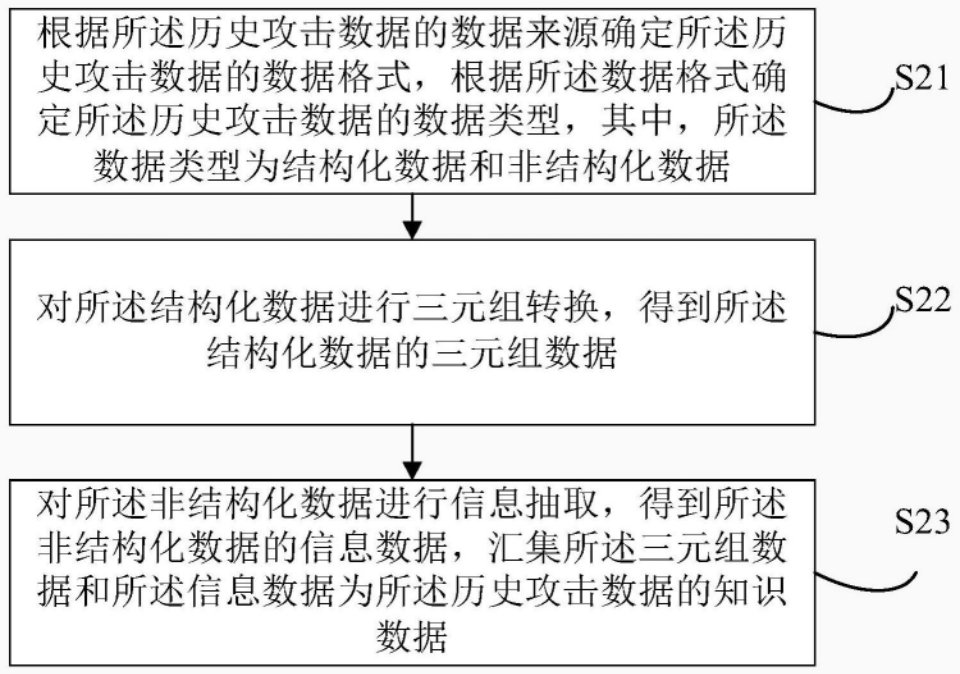

根据所述历史攻击数据的数据来源确定所述历史攻击数据的数据格式，根据所述数据格式确定所述历史攻击数据的数据类型，其中，所述数据类型为结构化数据和非结构化数据
S21
对所述结构化数据进行三元组转换，得到所述结构化数据的三元组数据
S22
对所述非结构化数据进行信息抽取，得到所述非结构化数据的信息数据，汇集所述三元组数据和所述信息数据为所述历史攻击数据的知识数据
S23

图2

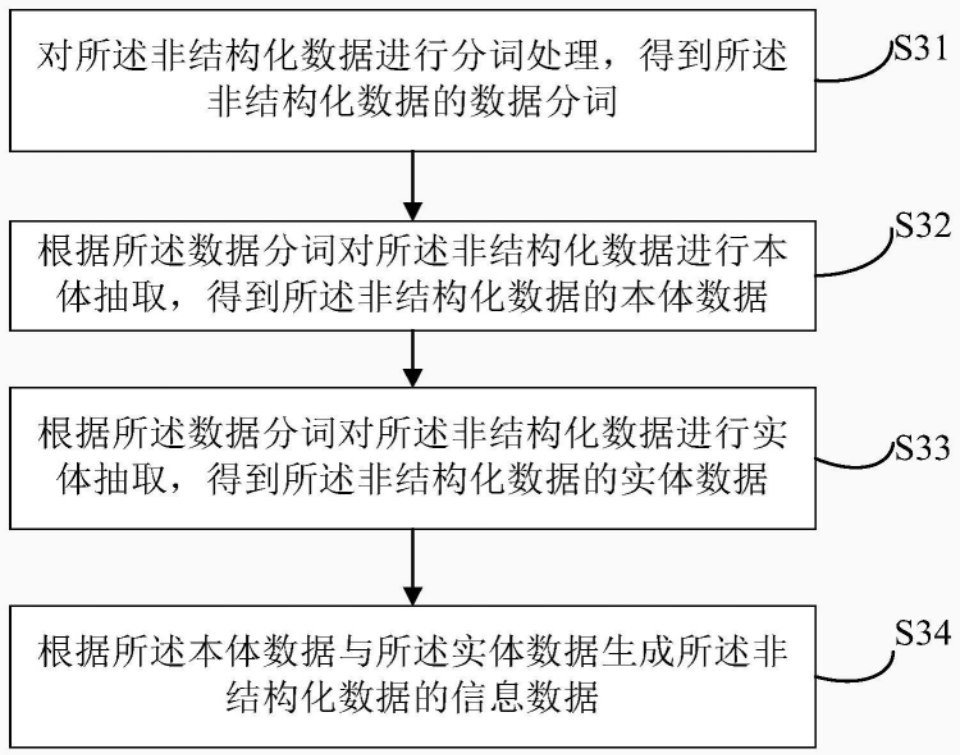

对所述非结构化数据进行分词处理，得到所述非结构化数据的数据分词
S31
根据所述数据分词对所述非结构化数据进行本体抽取，得到所述非结构化数据的本体数据
S32
根据所述数据分词对所述非结构化数据进行实体抽取，得到所述非结构化数据的实体数据
S33
根据所述本体数据与所述实体数据生成所述非结构化数据的信息数据
S34

图3

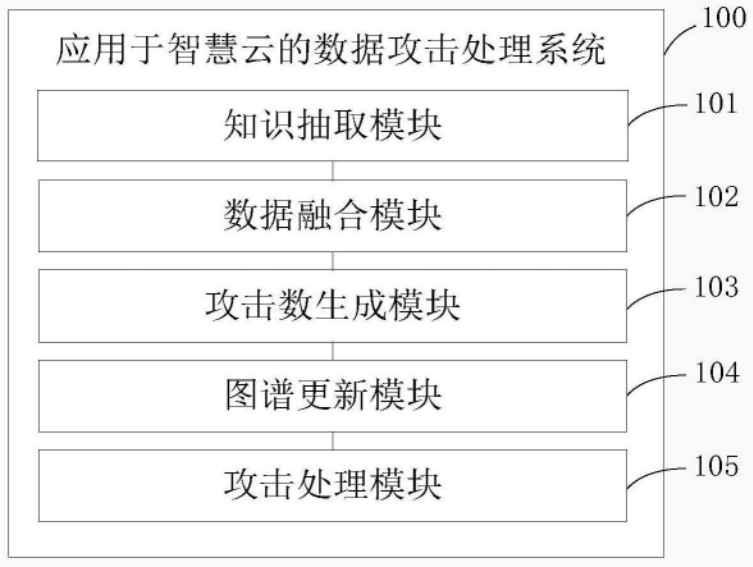

应用于智慧云的数据攻击处理系统
100
知识抽取模块
101
数据融合模块
102
攻击数生成模块
103
图谱更新模块
104
攻击处理模块
105

图4