

一种核身方法及装置

技术领域

[01] 本申请涉及信息处理技术领域，特别涉及一种核身方法及装置、计算设备和存储介质。

5 背景技术

[02] 对于金融类的应用程序，基于盗用风险防控的原因，可能会对某些场景下的部分用户输出某种核实身份的方式，用来核实用户的真实身份，降低用户的账号被盗用的风险。

- [03] 现有技术中，核身方式一般为一种单一固定的方式。这种核身方式存在两种弊端：
- 10 一是单一的核身方式可能存在被坏人突破的危险；二是单一核身对于风险低的用户可能存在不适用的情况，例如某个用户可能偏好人脸校验的核身方式，如果推送一个基于用户记忆的动态问卷进行核身，可能导致该用户的核身不能通过，从而影响用户体验。

发明内容

- [04] 有鉴于此，本说明书一个或多个实施例提供了一种核身方法及装置、计算设备和
- 15 存储介质，以解决现有技术中存在的技术缺陷。

[05] 本说明书一实施例公开了一种核身方法，所述核身方法包括：

[06] 获取用户的属性特征；

[07] 将用户的属性特征输入至训练好的核身安全模型，并根据输出的结果得到与该用户的属性特征相对应的第一核身产品以及第一核身产品的第一核身通过率；

- 20 [08] 将第一核身产品的任意两个生成第一核身产品组合，并根据第一核身产品的第一核身通过率计算所述第一核身产品组合的第二核身通过率，得到第二核身通过率大于第一阈值的所述第一核身产品组合中的第二核身产品；

[09] 将第二核身产品以及用户的属性特征输入至训练好的核身适用模型，得到所述第二核身产品的第三核身通过率；

- 25 [10] 根据第二核身产品的第三核身通过率计算得到所述第二核身产品组合的第四核身

通过率;

[11] 输出第四核身通过率大于第二阈值的第二核身产品组合。

[12] 本说明书一实施例公开了一种核身安全模型的生成方法, 所述方法包括:

5 [13] 获取核身安全样本和第一标签, 所述核身安全样本包括历史突破事件、该历史突破事件对应的核身产品和用户的历史属性特征, 所述第一标签为所述历史突破事件对应的核身产品核身通过或不通过;

[14] 建立核身安全模型, 将核身安全样本和第一标签输入至所述核身安全模型进行训练;

[15] 保存所述核身安全模型。

10 [16] 本说明书一实施例公开了一种核身适用模型的生成方法, 包括:

[17] 获取核身适用样本和第二标签, 所述核身适用样本包括历史非突破事件、该历史非突破事件对应的核身产品和用户的历史属性特征, 所述第二标签为所述历史非突破事件对应的核身产品核身通过或不通过;

15 [18] 建立核身适用模型, 将核身适用样本和第二标签输入至所述核身适用模型进行训练;

[19] 保存所述核身适用模型。

[20] 本说明书一实施例公开了一种核身装置, 所述装置包括:

[21] 属性特征获取模块, 被配置为获取用户的属性特征;

20 [22] 第一核身安全处理模块, 被配置为将用户的属性特征输入至训练好的核身安全模型, 并根据输出的结果得到与该用户的属性特征相对应的第一核身产品以及第一核身产品的第一核身通过率;

[23] 第二核身安全处理模块, 被配置为将第一核身产品的任意两个生成第一核身产品组合, 并根据第一核身产品的第一核身通过率计算所述第一核身产品组合的第二核身通过率, 得到第二核身通过率大于第一阈值的所述第一核身产品组合中的第二核身产品;

25 [24] 第一核身适用处理模块, 被配置为将第二核身产品以及用户的属性特征输入至训练好的核身适用模型, 得到所述第二核身产品的第三核身通过率;

[25] 第二核身适用处理模块, 被配置为根据第二核身产品的第三核身通过率计算得到

所述第二核身产品组合的第四核身通过率，并输出第四核身通过率大于第二阈值的第二核身产品组合。

[26] 本说明书一实施例公开了一种核身安全模型的生成装置，包括：

5 [27] 第一获取模块，被配置为获取核身安全样本和第一标签，所述核身安全样本包括历史突破事件、该历史突破事件对应的核身产品和用户的历史属性特征，所述第一标签为所述历史突破事件对应的核身产品核身通过或不通过；

[28] 第一训练模块，被配置为建立核身安全模型，将核身安全样本和第一标签输入至所述核身安全模型进行训练；

[29] 第一存储模块，被配置为保存所述核身安全模型。

10 [30] 本说明书一实施例公开了一种核身适用模型的生成装置，包括：

[31] 第二获取模块，被配置为获取核身适用样本和第二标签，所述核身适用样本包括历史非突破事件、该历史非突破事件对应的核身产品和用户的历史属性特征，所述第二标签为所述历史非突破事件对应的核身产品核身通过或不通过；

15 [32] 第二训练模块，被配置为建立核身适用模型，将核身适用样本和第二标签输入至所述核身适用模型进行训练；

[33] 第二存储模块，被配置为保存所述核身适用模型。

20 [34] 本说明书一实施例公开了一种计算设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机指令，所述处理器执行所述指令时实现该指令被处理器执行时实现如前所述核身方法或核身安全模型的生成方法或核身适用模型的生成方法的步骤。

[35] 本说明书一实施例公开了一种计算机可读存储介质，其存储有计算机指令，该指令被处理器执行时实现如前所述核身方法或核身安全模型的生成方法或核身适用模型的生成方法的步骤。

25 [36] 本说明书提供的核身方法及装置、计算设备和存储介质，在获取用户的属性特征后，依次输入训练好的核身安全模型和核身适用模型，并最终输出第四核身通过率大于第二阈值的任意两个第二核身产品生成的第二核身产品组合，从而提高了安全性，提升了用户体验。

附图说明

- [37] 图 1 是本说明书一个或多个实施例的计算设备的结构示意图;
- [38] 图 2 是本说明书一个或多个实施例的核身方法的流程示意图;
- [39] 图 3 是本说明书一个或多个实施例的核身方法的流程示意图;
- 5 [40] 图 4a 是本说明书一个或多个实施例的核身方法的流程示意图;
- [41] 图 4b 是本说明书一个或多个实施例的核身方法的核身通过率的结果展示图;
- [42] 图 5 是本说明书一个或多个实施例的核身安全模型的生成方法的流程示意图;
- [43] 图 6 是本说明书一个或多个实施例的核身适用模型的生成方法的流程示意图;
- [44] 图 7 是本说明书一个或多个实施例的核身装置的结构示意图;
- 10 [45] 图 8 是本说明书一个或多个实施例的核身安全模型的生成装置的结构示意图;
- [46] 图 9 是本说明书一个或多个实施例的核身适用模型的生成装置的结构示意图。

具体实施方式

- [47] 在下面的描述中阐述了很多具体细节以便于充分理解本申请。但是本申请能够以很多不同于在此描述的其它方式来实施，本领域技术人员可以在不违背本申请内涵的情况下做类似推广，因此本申请不受下面公开的具体实施的限制。
- 15

- [48] 在本说明书一个或多个实施例中使用的术语是仅仅出于描述特定实施例的目的，而非旨在限制本说明书一个或多个实施例。在本说明书一个或多个实施例和所附权利要求书中所使用的单数形式的“一种”、“所述”和“该”也旨在包括多数形式，除非上下文清楚地表示其他含义。还应当理解，本说明书一个或多个实施例中使用的术语“和/或”是指并包含一个或多个相关联的列出项目的任何或所有可能组合。
- 20

- [49] 应当理解，尽管在本说明书一个或多个实施例中可能采用术语第一、第二等来描述各种信息，但这些信息不应限于这些术语。这些术语仅用来将同一类型的信息彼此区分开。例如，在不脱离本说明书一个或多个实施例范围的情况下，第一也可以被称为第二，类似地，第二也可以被称为第一。取决于语境，如在此所使用的词语“如果”可以被解释成为“在……时”或“当……时”或“响应于确定”。
- 25

- [50] 首先，对本发明一个或多个实施例涉及的名词术语进行解释。

- [51] 核身：即核实身份的简称，为了保证用户账号的安全，在用户使用应用程序时进行的身份验证。例如金融类的应用程序，为了防止盗用风险，在用户登录或者支付时进行的身份验证。
- [52] 核身产品：用于进行核实身份的产品。例如基于信息的核身产品包括：隐私问题、预留信息等；基于 PKI（Public Key Infrastructure，公钥基础设施）技术的核身产品：支付盾、数字证书等；基于 OTP（One-time Password，动态口令）技术的核身产品：手机宝令、一键确认、动态短信校验码等；基于生物特征的核身产品：指纹等；基于行为特征的核身产品：笔迹等。
- 10 [53] 马尔可夫链蒙特卡罗（Markov Chain Monte Carlo，MCMC）：为贝叶斯理论框架下，通过计算机进行模拟的蒙特卡罗模型。该模型的算法为将马尔科夫(Markov)过程引入到蒙特卡罗（Monte Carlo）模拟中，实现抽样分布随模拟的进行而改变的动态模拟，弥补了传统的蒙特卡罗积分只能静态模拟的缺陷。
- 15 [54] 随机森林（Random Decision Forests）：在机器学习中，随机森林是一个包含多个分类树的分类器。把分类树组合成随机森林，即在变量（列）的使用和数据（行）的使用上进行随机化，生成很多分类树，再汇总分类树的结果。
- [55] 突破事件：应当被拦截却未拦截的事件，例如被盗用事件。
- [56] 在本说明书中，提供了一种核身方法及装置、计算设备和计算机可读存储介质，在下面的实施例中逐一进行详细说明。
- 20 [57] 图 1 是示出了根据本说明书一实施例的计算设备 100 的结构框图。该计算设备 100 的部件包括但不限于存储器 110 和处理器 120。处理器 120 与存储器 110 通过总线 130 相连接，数据库 150 用于保存数据，网络 160 用于接收数据库 150 存储的数据。
- 25 [58] 计算设备 100 还包括接入设备 140，接入设备 140 使得计算设备 100 能够经由一个或多个网络 160 通信。这些网络的示例包括公用交换电话网（PSTN）、局域网（LAN）、广域网（WAN）、个域网（PAN）或诸如因特网的通信网络的组合。接入设备 140 可以包括有线或无线的任何类型的网络接口（例如，网络接口卡（NIC））中的一个或多个，诸如 IEEE802.11 无线局域网（WLAN）无线接口、全球微波互联接入（Wi-MAX）接口、以太网接口、通用串行总线（USB）接口、蜂窝网络接口、蓝牙接口、近场通信（NFC）接口，等等。
- [59] 在本说明书的一个实施例中，计算设备 100 的上述以及图 1 中未示出的其他部件

也可以彼此相连接，例如通过总线。应当理解，图 1 所示的计算设备结构框图仅仅是出于示例的目的，而不是对本说明书范围的限制。本领域技术人员可以根据需要，增添或替换其他部件。

5 [60] 计算设备 100 可以是任何类型的静止或移动计算设备，包括移动计算机或移动计算设备（例如，平板计算机、个人数字助理、膝上型计算机、笔记本计算机、上网本等）、移动电话（例如，智能手机）、可佩戴的计算设备（例如，智能手表、智能眼镜等）或其他类型的移动设备，或者诸如台式计算机或 PC 的静止计算设备。计算设备 100 还可以是移动式或静止式的服务器。

10 [61] 其中，处理器 120 可以执行图 2 所示方法中的步骤。图 2 是示出了根据本说明书一个或多个实施例的核身方法的示意性流程图，包括步骤 202 至步骤 212：

[62] 202、获取用户的属性特征。

15 [63] 具体地，用户的属性特征可以包括多种，例如用户信息，包括性别、年龄、城市、注册时长、绑卡、等级等；用户层次，包括学历、职业、财富、信用、消费、房产、车辆等；用户行为，包括历史核身、购买行为、安全意识等；用户偏好，包括网络偏好、无线偏好等；用户设备，包括品牌、通讯录、版本信息、系统等；用户场景，包括家庭、公共场合、公司、酒店、打车、扫码、转账等。

[64] 204、将用户的属性特征输入至训练好的核身安全模型，并根据输出的结果得到与该用户的属性特征相对应的第一核身产品以及第一核身产品的第一核身通过率。

20 [65] 其中，核身安全模型在训练时，使用的训练样本均为历史突破事件。所以，将用户的属性特征输入至核身安全模型，其实质是假定该用户进行的操作事件为突破事件，然后得到该用户对应的第一核身产品的第一核身通过率，以确定适用于该用户的核身产品。

[66] 具体地，参见图 3，本步骤 204 包括：

25 [67] 2042、将用户的属性特征输入至训练好的核身安全模型，得到与该用户的属性特征相对应的第一核身产品的第五核身通过率。

[68] 2044、统计目标时间范围内的与该用户的属性特征相对应的第一核身产品的第六核身通过率。

[69] 具体地，步骤 2044 中统计目标时间范围内的与该用户的属性特征相对应的第一核

身产品的第六核身通过率，包括：

[70] 将目标时间范围内的风险分为风险群组；

[71] 统计每个风险群组的与所述第一核身产品对应的核身通过率，将各个风险群组的所述核身通过率加权，得到第一核身产品的第六核身通过率。

5 [72] 目标时间范围可以根据需求设定，例如设定目标时间范围为最近十天。

[73] 第五核身通过率和第六核身通过率不同，第五核身通过率是表征第一核身产品的长期风险，第六核身通过率是表征第一核身产品的短期风险。设置第五核身通过率和第六核身通过率的原因是由于现有的数据体系难以覆盖所有的事件，例如用户与黑产分子在线下沟通并深度被骗，用户主动泄露信息的场景。通过统计第五核身通过率和第六核
10 身通过率，以提高核身安全模型的性能。

[74] 其中，第一核身产品的第六核身通过率通过统计得到。在统计第一核身产品的第六核身通过率时，需要根据不同的第一核身产品的容易被突破的因素，设计第六核身通过率所对应的风险，并根据风险得分计算得到第六核身通过率。表1中列举了几种不同的第一核身产品的突破因素，表2中列举了根据不同的第一核身产品的突破因素设计的
15 风险统计。表2中，将风险按照行为、设备、环境、用户和交易进行分类，再统计每种第一核身产品的所对应的风险得分，进而统计出第一核身产品的第六核身通过率。

表 1

第一核身产品	突破因素
KBA	1) 半动态信息, 且输出信息不易传送, 不易发生在电信欺诈场景; 2) 信息泄漏足够严重的情况下, 扫号、洗号会严重;
扫码鉴权	1) 半动态信息, 不易发生扫号案件, 发生的均为前期换绑了手机; 2) 对于扫码, 重视程度不够, 易发生欺诈;
安保问题	1) 静态信息, 易发生洗号案件; 2) 对于手机丢失、电信欺诈防控效果佳;
账户电话认证	1) 洗号案件以前期换绑手机为主; 2) 对于打电话, 重视程度不够, 欺诈案件居多;
绑定手机	1) 手机木马, 洗号换绑手机案件为主; 2) 动态信息, 社工、钓鱼案件相对不多; 3) 目前对短信重视程度增加, 突破率下降;
可信手机	类同绑定手机
人脸	1) 安全级别高, 体验不好; 2) 静态信息, 高级盗用者容易掌握并模拟; 3) 有可能发生欺诈案件;

表 2

	行为	设备	环境	用户	交易
风险	近期是否泄漏 SMS/电话	Umid/Apdid 关联认证方式	扫码鉴权地与支付地非同市	是否同证件近期出现绑卡	操作渠道
	近期是否泄漏密码	Umid/Apdid 关联同认证方式账户个数	IP 关联认证通过方式	关联账户个数	场景
	近期是否泄漏安保问题	Umid/Apdid 关联同认证方式通过/未通过用户数	IP 关联同认证方式账户个数	是否稳定使用未越狱 IOS 设备	核身方式类型 (动态、静态、半动态)
	近期是否泄漏证件信息	Umid/Apdid 关联同认证方式通过率	IP 关联同认证方式通过/未通过用户数	是否稳定使用未 PC 设备	核身方式相似度
	近期是否泄漏快捷卡信息		IP 关联同认证方式通过率	年龄	认证方式突破案件数
	近期是否有 KBA 通过记录		异省	性别	风险类型
	近期是否有查询交易记录			账户活跃程度	
	近期是否浏览账户主页/财富页面信息				
	前期未通过认证信息列表				

[75] 2046、将第一核身产品的第五核身通过率和第六核身通过率进行比较，取较大值作为第一核身产品的第一核身通过率。

[76] 例如在一个具体实施方案中，通过计算第五核身通过率和第六核身通过率，得到

5 第一核身产品的第五核身通过率为 0.8，第六核身通过率为 0.9。那么，第六核身通过率

要比第五核身通过率大。将第一核身产品的第六核身通过率 0.9 作为输出的第一核身产品的第一核身通过率进行后续的计算，从而可以避免近期较大的批量风险。

- [77] 206、将第一核身产品的任意两个生成第一核身产品组合，并根据第一核身产品的第一核身通过率计算所述第一核身产品组合的第二核身通过率，将第二核身通过率大于第一阈值的所述第一核身产品组合作为第二核身产品组合，得到第二核身产品组合中的第二核身产品。

[78] 需要注意的是，第二核身产品组合为第一核身产品组合的子集，在理想情况下，第一核身产品组合和第二核身产品组合相同。第二核身产品为第一核身产品的子集，在理想情况下，第一核身产品和第二核身产品的个数相等。

- 10 [79] 具体地，所述第一核身产品组合的第二核身通过率通过以下公式计算：

$$P(A+B)=1-(1-P(A))*(1-P(B))$$

[80] 其中， $P(A+B)$ 为第一核身产品组合的第二核身通过率；

[81] $P(A)$ 、 $P(B)$ 分别为第一核身产品 A 和 B 的第一核身通过率。

- [82] 本步骤中，将第二核身通过率大于第一阈值的第一核身产品组合作为第二核身产品组合输出，与输出单个第一核身产品的技术方案相比，可以提高核身的安全性。

[83] 本步骤中，第一阈值可以根据实际需求而设置，例如设置第一阈值为 0.8，取第二核身通过率大于 0.8 的第一核身产品组合中的第一核身产品作为第二核身产品。

[84] 可选地，在得到第二核身产品组合中的第二核身产品后，还包括：

- [85] 根据预先设置的安全管控策略，对得到的第二核身产品进行筛选，并将筛选后的第二核身产品输入至训练好的核身适用模型。

[86] 安全管控策略为预先设置，以过滤掉对该用户不匹配的第二核身产品。

[87] 例如，用户的移动设备较旧，没有前置摄像头，那么人脸验证的核身产品对于该用户便不适用。

- [88] 208、将第二核身产品以及用户的属性特征输入至训练好的核身适用模型，得到第二核身产品的第三核身通过率。

[89] 本实施例中，将第二核身产品输入至核身适用模型，可以确定第二核身产品是否对于用户适用。

- [90] 本步骤中,核身适用模型可以为多种,例如随机森林算法模型,然后得到 F1 Score、KS、AUC 的分值,进而得出第二核身产品组合中的第二核身产品的第三核身通过率。
- [91] 210、根据第二核身产品的第三核身通过率计算得到所述第二核身产品组合的第四核身通过率。
- 5 [92] 本步骤中,在得到第二核身产品的第三核身通过率后,通过公式计算得到第二核身产品组合的第四核身通过率。
- [93] 具体地,第二核身产品组合的第四核身通过率通过以下公式计算:
- $$P(C+D)=1-(1-P(C))*(1-P(D))$$
- [94] 其中, $P(C+D)$ 为第二核身产品组合的第四核身通过率;
- 10 [95] $P(C)$ 、 $P(D)$ 分别为第二核身产品的第三核身通过率。
- [96] 212、输出第四核身通过率大于第二阈值的第二核身产品组合。
- [97] 本步骤中,第二阈值可以根据实际需求而设置,例如设置第二阈值为 0.9,输出第四核身通过率大于 0.9 的第二核身产品组合。
- [98] 例如第二核身产品包括人脸验证和短信验证,人脸验证的第三核身通过率为 0.6,
- 15 短信验证的第三核身通过率 0.8,可以计算得到人脸验证+短信验证的第二核身产品组合的第四核身通过率为 0.92。那么,输出人脸验证+短信验证的第二核身产品组合。
- [99] 本说明书提供的核身方法,在获取用户的属性特征后,依次输入训练好的核身安全模型和核身适用模型,并最终输出第四核身通过率大于第二阈值的任意两个第二核身产品生成的第二核身产品组合,从而提高了安全性,提升了用户体验。
- 20 [100] 并且,通过核身安全模型,在假定用户的事件为安全事件的情形下,获取与该用户对应的第一核身产品的第一核身通过率,从而可以得到适用于该用户的核身安全性较高的第一核身产品以及第一核身产品组合。
- [101] 通过核身适用模型,在假定用户的事件为非安全事件的情况下,获取与该用户对应的第二核身产品的第三核身通过率,从而可以得到适用于该用户的核身适用性较高的
- 25 第二核身产品以及第二核身产品组合。
- [102] 为了便于理解本说明书记载的技术方案,下面以一个具体实例进行说明。参见图 4a 和图 4b。其中,图 4b 示出了本实施例的核身方法的核身通过率的结果展示。
- [103] 参见图 4a,本实施例的核身方法包括:

- [104] 402、设置安全管控策略。
- [105] 404、获取用户的属性特征。
- [106] 406、通过核身安全模型获取第一核身产品的第一核身通过率。
- [107] 本实施例中，该用户对应的第一核身产品包括：人脸、短信、基于知识的身份验证（KBA）和安保。
- 5 [108] 参见图 4b，经过计算，得到人脸的第一核身通过率为 0.7，短信的第一核身通过率为 0.6，KBA 的第一核身通过率为 0.5，安保的第一核身通过率为 0.4。
- [109] 408、计算第一核身产品组合的第二核身通过率。
- [110] 参见上述实施例，根据第一核身产品的第一核身通过率计算所述第一核身产品组合的第二核身通过率。
- 10 [111] 参见图 4b，经过计算，得到人脸+短信的第二核身通过率为 0.88，KBA+安保的第二核身通过率为 0.7，短信+安保的第二核身通过率为 0.76，KBA+短信的第二核身通过率为 0.8。
- [112] 410、将第二核身通过率大于第一阈值的所述第一核身产品组合作为第二核身产品组合，得到第二核身产品组合中的第二核身产品。
- 15 [113] 本实施例中，选择第一阈值为 0.7，将第二核身通过率大于 0.7 的第一核身产品组合作为第二核身产品组合，得到第二核身产品组合中的第二核身产品。
- [114] 根据上述结果，第二核身产品组合包括：人脸+短信、短信+安保、KBA+短信。其中的第二核身产品包括：人脸、短信、KBA、安保。
- 20 [115] 412、通过核身适用模型得到第二核身产品的第三核身通过率。
- [116] 步骤 410 得到的第二核身产品还需要经过步骤 402 的安全管控策略的过滤，然后得到输入至核身适用模型的第二核身产品。
- [117] 参见图 4b，经过计算，得到人脸的第三核身通过率为 0.5，短信的第三核身通过率为 0.5，KBA 的第三核身通过率为 0.6，安保的第三核身通过率为 0.7。
- 25 [118] 414、计算得到核第二身产品组合的第四核身通过率，输出第四核身通过率大于第二阈值的第二核身产品组合。
- [119] 参见图 4b，根据上述第二核身产品的第三核身通过率，计算人脸+短信的第四核身

通过率为 0.75、短信+安保的第四核身通过率为 0.85、KBA+短信的第四核身通过率为 0.8。

[120] 本实施例中，选择第二阈值为 0.75。输出第四核身通过率大于 0.75 的第二核身产品组合包括：短信+安保的组合以及 KBA+短信的组合，以供用户选择。

- 5 [121] 如果用户选择短信+安保的第二核身产品组合，便在用户使用应用程序的过程中，加载短信+安保的第二核身产品组合进行核身。

[122] 本说明书一实施例还公开了一种核身安全模型的生成方法，参见图 5，包括：

- 10 [123] 502、获取核身安全样本和第一标签，所述核身安全样本包括历史突破事件、该历史突破事件对应的核身产品和用户的历史属性特征，所述第一标签为所述历史突破事件对应的核身产品核身通过或不通过。

[124] 由本步骤可见，核身安全性模型的样本，包括黑样本，即核身通过的历史突破事件；以及白样本，即核身不通过的历史突破事件。

[125] 504、建立核身安全模型，将核身安全样本和第一标签输入至所述核身安全模型进行训练；

- 15 [126] 506、保存所述核身安全模型。

[127] 实际使用时，单个核身产品的安全性的衡量，需要解决样本量小的问题：一个问题黑白样本都少，不足以支撑使用常规的机器学习模型方法；另一个问题是黑样本少，数据量级不足以支撑使用常规的机器学习模型方法。那么，本实施例中，使用安全模型为马尔可夫链蒙特卡罗 MCMC 模型，可以解决样本量小的问题，具体处理过程如下：

- 20 [128] 1) 对已有核身安全样本进行分析，得到样本分布。

[129] 假定五个参数（回归系数）相互独立，并服从均值为 0、标准差为 X 的先验正态分布。

[130] 2) 定义非标准的对数后验发布概率密度（即先验函数与拟然函数相乘）。

[131] 3) 产生随机数据，模拟核身安全模型的后验分布。

- 25 [132] 也就是说，MCMC 模型的样本要通过从后验分布中进行取样而获得。

[133] 4) 调整 proposal（建议分布）的 scale 参数（即随机游走 MH 算法的方差）来获得一个更合理的接受概率。

- [134] 改变 scale 参数, 参数值越大, 那么马尔可夫链状态转移就越明显 (样本自相关性减弱), 但同时接受概率就越低, 所以这是一个需要兼顾两者的选择。
- [135] 5) 找到了适合接受概率的 scale 值, 按此参数增加模拟次数, 然后观察结果是否收敛, 在结果收敛的情况下, 获得模拟出来的样本。
- 5 [136] 模拟出来的样本, 其数据量级要比初始的核身安全样本的数据量级大, 从而可以满足 MCMC 模型训练的需要。
- [137] 6) 基于模拟出来的样本进行训练, 得到初始核身安全性模型。
- [138] 在后期, 依据上线后不断新增的样本, 对该核身安全模型进行优化。
- [139] 通过本实施例生成的核身安全模型, 在获知突破事件后, 可以在输入获知该突破事件对应的第一核身产品以及第一核身产品的第一核身通过率, 从而确定与适用于判断该突破事件的核身安全性较高的核身产品。
- 10 [140] 并且, 通过核身安全模型, 可以在假定用户的事件为安全事件的情形下, 获取与该用户对应的第一核身产品的第一核身通过率, 从而可以得到适用于该用户的核身安全性较高的第一核身产品以及第一核身产品组合。
- 15 [141] 本说明书一实施例还公开了一种核身适用模型的生成方法, 参见图 6, 包括:
- [142] 602、获取核身适用样本和第二标签, 所述核身适用样本包括历史非突破事件、该历史非突破事件对应的核身产品和用户的历史属性特征, 所述第二标签为所述历史非突破事件对应的核身产品核身通过或不通过。
- [143] 由本步骤可见, 核身适用性模型的样本, 包括黑样本, 即核身通过的历史非突破事件; 以及白样本, 即核身不通过的历史非突破事件。
- 20 [144] 604、建立核身适用模型, 将核身适用样本和第二标签输入至所述核身适用模型进行训练。
- [145] 606、保存所述核身适用模型。
- [146] 本实施例中, 生成的核身适用性模型可以用于判断输入事件所对应的核身产品与用户的历史属性特征是否相匹配。
- 25 [147] 本实施例中, 核身适用模型可以为随机森林算法模型。
- [148] 在机器学习中, 随机森林算法由许多的决策树组成, 因为这些决策树的形成采用

了随机的方法，因此也叫做随机决策树。随机森林算法模型中的树之间是没有关联的。当测试数据进入随机森林算法模型时，其实就是让每一颗决策树进行分类，最后取所有决策树中分类结果最多的那类为最终的结果。因此随机森林算法模型是一个包含多个决策树的分类器，并且其输出的类别是由个别树输出的类别的众数而定。随机森林算法模型既可以处理属性为离散值的量，如 ID3 算法，也可以处理属性为连续值的量，比如 C4.5 算法。

[149] 随机森林算法模型的构造过程如下：

[150] 1) 假如有 N 个样本，则有放回的随机选择 N 个样本（每次随机选择一个样本，然后返回继续选择）。这选择好了的 N 个样本用来训练一个决策树，作为决策树根节点处的样本。

[151] 2) 当每个样本有 M 个属性时，在决策树的每个节点需要分裂时，随机从这 M 个属性中选取 m 个属性，满足条件 $m \ll M$ 。然后从这 m 个属性中采用某种策略（比如说信息增益）来选择 1 个属性作为该节点的分裂属性。

[152] 3) 决策树形成过程中每个节点都要按照步骤 2 来分裂（很容易理解，如果下一次该节点选出来的那一个属性是刚刚其父节点分裂时用过的属性，则该节点已经达到了叶子节点，无须继续分裂了）。一直到不能够再分裂为止。注意整个决策树形成过程中没有进行剪枝。

[153] 4) 按照步骤 1~3 建立大量的决策树，这样就构成了随机森林模型。

[154] 本实施例中，生成的核身适用性模型可以用于判断输入事件所对应的核身产品与用户的历史属性特征是否相匹配。

[155] 通过核身适用模型，在假定用户的事件为非安全事件的情况下，获取与该用户对应的第二核身产品的第三核身通过率，从而可以得到适用于该用户的核身适用性较高的第二核身产品以及第二核身产品组合。

[156] 本说明书一实施例还提供一种核身装置，参见图 7，包括：

[157] 属性特征获取模块 702，被配置为获取用户的属性特征；

[158] 第一核身安全处理模块 704，被配置为将用户的属性特征输入至训练好的核身安全模型，并根据输出的结果得到与该用户的属性特征相对应的第一核身产品以及第一核身产品的第一核身通过率；

- [159] 第二核身安全处理模块 706, 被配置为将第一核身产品的任意两个生成第一核身产品组合, 并根据第一核身产品的第一核身通过率计算所述第一核身产品组合的第二核身通过率, 将第二核身通过率大于第一阈值的所述第一核身产品组合作为第二核身产品组合, 得到第二核身产品组合中的第二核身产品;
- 5 [160] 第一核身适用处理模块 708, 被配置为将第二核身产品以及用户的属性特征输入至训练好的核身适用模型, 得到第二核身产品的第三核身通过率;
- [161] 第二核身适用处理模块 710, 被配置为根据第二核身产品的第三核身通过率计算得到所述第二核身产品组合的第四核身通过率, 并输出第四核身通过率大于第二阈值的第二核身产品组合。
- 10 [162] 可选地, 所述第一核身安全处理模块 704 包括:
- [163] 第五核身通过率获取模块 7042, 被配置为将用户的属性特征输入至训练好的核身安全模型, 得到与该用户的属性特征相对应的第一核身产品的第五核身通过率;
- [164] 第六核身通过率获取模块 7044, 被配置为统计目标时间范围内的与该用户的属性特征相对应的第一核身产品的第六核身通过率;
- 15 [165] 第一核身通过率确定模块 7046, 被配置为将第一核身产品的第五核身通过率和第六核身通过率进行比较, 取较大值作为第一核身产品的第一核身通过率。
- [166] 可选地, 第六核身通过率获取模块 7044 被配置为将目标时间范围内的风险分为风险群组; 统计每个风险群组的与所述第一核身产品对应的核身通过率, 将各个风险群组的所述核身通过率加权, 得到第一核身产品的第六核身通过率。
- 20 [167] 可选地, 在所述第二核身安全处理模块 706 得到第二核身产品组合中的第二核身产品后, 还包括:
- [168] 筛选模块, 被配置为根据预先设置的安全管控策略, 对得到的第二核身产品进行筛选, 并将筛选后的第二核身产品输入至训练好的核身适用模型。
- [169] 可选地, 所述第二核身安全处理模块 706 计算所述第一核身产品组合的第二核身
- 25 通过率通过以下公式:

$$P(A+B)=1-(1-P(A))*(1-P(B))$$

[170] 其中, $P(A+B)$ 为第一核身产品组合的第二核身通过率;

[171] $P(A)$ 、 $P(B)$ 分别为第一核身产品的第一核身通过率。

[172] 所述第二核身适用处理模块 710 计算第二核身产品组合的第四核身通过率通过以下公式:

$$P(C+D)=1-(1-P(C))*(1-P(D))$$

[173] 其中, $P(C+D)$ 为第二核身产品组合的第四核身通过率;

5 [174] $P(C)$ 、 $P(D)$ 分别为第二核身产品的第三核身通过率。

[175] 本说明书提供的核身装置, 在获取用户的属性特征后, 依次输入训练好的核身安全模型和核身适用模型, 并最终输出第四核身通过率大于第二阈值的任意两个第二核身产品生成的第二核身产品组合, 从而提高了安全性, 提升了用户体验。

[176] 本说明书一实施例还提供一种核身安全模型的生成装置, 参见图 8, 包括:

10 [177] 第一获取模块 802, 被配置为获取核身安全样本和第一标签, 所述核身安全样本包括历史突破事件、该历史突破事件对应的核身产品和用户的历史属性特征, 所述第一标签为所述历史突破事件对应的核身产品核身通过或不通过;

[178] 第一训练模块 804, 被配置为建立核身安全模型, 将核身安全样本和第一标签输入至所述核身安全模型进行训练;

15 [179] 第一存储模块 806, 被配置为保存所述核身安全模型。

[180] 上述为本实施例的一种核身安全模型的生成装置的示意性方案。需要说明的是, 该生成装置的技术方案与上述的核身安全模型的生成方法的技术方案属于同一构思, 该生成装置的技术方案未详细描述的细节内容, 均可以参见上述核身安全模型的生成方法的技术方案的描述。

20 [181] 本说明书一实施例公开了一种核身适用模型的生成装置, 参见图 9, 包括:

[182] 第二获取模块 902, 被配置为获取核身适用样本和第二标签, 核身适用样本包括历史非突破事件、该历史非突破事件对应的核身产品和用户的历史属性特征, 所述第二标签为所述历史非突破事件对应的核身产品核身通过或不通过;

25 [183] 第二训练模块 904, 被配置为建立核身适用模型, 将核身适用样本和第二标签输入至所述核身适用模型进行训练;

[184] 第二存储模块 906, 被配置为保存所述核身适用模型。

[185] 上述为本实施例的一种核身适用模型的生成装置的示意性方案。需要说明的是, 该生成装置的技术方案与上述的核身适用模型的生成方法的技术方案属于同一构思, 该

生成装置的技术方案未详细描述的细节内容，均可以参见上述核身适用模型的生成方法的技术方案的描述。

5 [186] 本说明书一实施例还提供一种计算设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机指令，所述处理器执行所述指令时实现该指令被处理器执行时实现如前所述核身方法或核身安全模型的生成方法或核身适用模型的生成方法的步骤。

[187] 本说明书一实施例还提供一种计算机可读存储介质，其存储有计算机指令，该指令被处理器执行时实现如前所述核身方法或核身安全模型的生成方法或核身适用模型的生成方法的步骤。

10 [188] 上述为本实施例的一种计算机可读存储介质的示意性方案。需要说明的是，该存储介质的技术方案与上述的核身方法或核身安全模型的生成方法或核身适用模型的生成方法的技术方案属于同一构思，存储介质的技术方案未详细描述的细节内容，均可以参见上述核身方法或核身安全模型的生成方法或核身适用模型的生成方法的技术方案的描述。

15 [189] 所述计算机指令包括计算机程序代码，所述计算机程序代码可以为源代码形式、对象代码形式、可执行文件或某些中间形式等。所述计算机可读介质可以包括：能够携带所述计算机程序代码的任何实体或装置、记录介质、U 盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器（ROM，Read-Only Memory）、随机存取存储器（RAM，Random Access Memory）、电载波信号、电信信号以及软件分发介质等。需要说明的是，
20 所述计算机可读介质包含的内容可以根据司法管辖区内立法和专利实践的要求进行适当的增减，例如在某些司法管辖区，根据立法和专利实践，计算机可读介质不包括电载波信号和电信信号。

[190] 需要说明的是，对于前述的各方法实施例，为了简便描述，故将其都表述为一系列的动作组合，但是本领域技术人员应该知悉，本申请并不受所描述的动作顺序的限制，
25 因为依据本申请，某些步骤可以采用其它顺序或者同时进行。其次，本领域技术人员也应该知悉，说明书中所描述的实施例均属于优选实施例，所涉及的动作和模块并不一定都是本申请所必须的。

[191] 在上述实施例中，对各个实施例的描述都各有侧重，某个实施例中没有详述的部分，可以参见其它实施例的相关描述。

[192] 以上公开的本申请优选实施例只是用于帮助阐述本申请。可选实施例并没有详尽叙述所有的细节，也不限制该发明仅为所述的具体实施方式。显然，根据本说明书的内容，可作很多的修改和变化。本说明书选取并具体描述这些实施例，是为了更好地解释本申请的原理和实际应用，从而使所属技术领域技术人员能很好地理解和利用本申请。

5 本申请仅受权利要求书及其全部范围和等效物的限制。

权利要求书

- 1、一种核身方法，其特征在于，所述核身方法包括：
- 获取用户的属性特征；
- 将用户的属性特征输入至训练好的核身安全模型，并根据输出的结果得到与该用户
- 5 的属性特征相对应的第一核身产品以及第一核身产品的第一核身通过率；
- 将第一核身产品的任意两个生成第一核身产品组合，并根据第一核身产品的第一核身通过率计算所述第一核身产品组合的第二核身通过率，将第二核身通过率大于第一阈值的所述第一核身产品组合作为第二核身产品组合，得到第二核身产品组合中的第二核身产品；
- 10 将第二核身产品以及用户的属性特征输入至训练好的核身适用模型，得到所述第二核身产品的第三核身通过率；
- 根据第二核身产品的第三核身通过率计算得到所述第二核身产品组合的第四核身通过率；
- 输出第四核身通过率大于第二阈值的第二核身产品组合。
- 15 2、如权利要求 1 所述的方法，其特征在于，将用户的属性特征输入至训练好的核身安全模型，并根据输出的结果得到与该用户的属性特征相对应的第一核身产品的第一核身通过率，包括：
- 将用户的属性特征输入至训练好的核身安全模型，得到与该用户的属性特征相对应
- 的第一核身产品的第五核身通过率；
- 20 统计目标时间范围内的与该用户的属性特征相对应的第一核身产品的第六核身通过率；
- 将第一核身产品的第五核身通过率和第六核身通过率进行比较，取较大值作为第一核身产品的第一核身通过率。
- 3、如权利要求 2 所述的方法，其特征在于，统计目标时间范围内的与该用户的属
- 25 性特征相对应的第一核身产品的第六核身通过率，包括：
- 将目标时间范围内的风险分为风险群组；
- 统计每个风险群组的与所述第一核身产品对应的核身通过率，将各个风险群组的所述核身通过率加权，得到第一核身产品的第六核身通过率。
- 4、如权利要求 1 所述的方法，其特征在于，在得到第二核身产品组合中的第二核
- 30 身产品后，还包括：
- 根据预先设置的安全管控策略，对得到的第二核身产品进行筛选，并将筛选后的第

二核身产品输入至训练好的核身适用模型。

5、如权利要求 1 所述的方法，其特征在于，所述第一核身产品组合的第二核身通过率通过以下公式计算：

$$P(A+B)=1-(1-P(A))*(1-P(B))$$

5 其中， $P(A+B)$ 为第一核身产品组合的第二核身通过率；

$P(A)$ 、 $P(B)$ 分别为第一核身产品 A 和 B 的第一核身通过率。

6、如权利要求 1 所述的方法，其特征在于，所述第二核身产品组合的第四核身通过率通过以下公式计算：

$$P(C+D)=1-(1-P(C))*(1-P(D))$$

10 其中， $P(C+D)$ 为第二核身产品组合的第四核身通过率；

$P(C)$ 、 $P(D)$ 分别为第二核身产品的第三核身通过率。

7、一种核身安全模型的生成方法，其特征在于，所述方法包括：

获取核身安全样本和第一标签，所述核身安全样本包括历史突破事件、该历史突破事件对应的核身产品和用户的历史属性特征，所述第一标签为所述历史突破事件对应的

15 核身产品核身通过或不通过；

建立核身安全模型，将核身安全样本和第一标签输入至所述核身安全模型进行训练；
保存所述核身安全模型。

8、如权利要求 7 所述的方法，其特征在于，所述核身安全模型为马尔可夫链蒙特卡罗 MCMC 模型。

20 9、一种核身适用模型的生成方法，其特征在于，包括：

获取核身适用样本和第二标签，所述核身适用样本包括历史非突破事件、该历史非突破事件对应的核身产品和用户的历史属性特征，所述第二标签为所述历史非突破事件对应的核身产品核身通过或不通过；

25 建立核身适用模型，将核身适用样本和第二标签输入至所述核身适用模型进行训练；
保存所述核身适用模型。

10、如权利要求 9 所述的方法，其特征在于，所述核身适用模型为随机森林算法模型。

11、一种核身装置，其特征在于，所述装置包括：

属性特征获取模块，被配置为获取用户的属性特征；

30 第一核身安全处理模块，被配置为将用户的属性特征输入至训练好的核身安全模型，并根据输出的结果得到与该用户的属性特征相对应的第一核身产品以及第一核身产品

的第一核身通过率;

第二核身安全处理模块,被配置为将第一核身产品的任意两个生成第一核身产品组合,并根据第一核身产品的第一核身通过率计算所述第一核身产品组合的第二核身通过率,将第二核身通过率大于第一阈值的所述第一核身产品组合作为第二核身产品组合,

5 得到第二核身产品组合中的第二核身产品;

第一核身适用处理模块,被配置为将第二核身产品以及用户的属性特征输入至训练好的核身适用模型,得到所述第二核身产品的第三核身通过率;

第二核身适用处理模块,被配置为根据第二核身产品的第三核身通过率计算得到所述第二核身产品组合的第四核身通过率,并输出第四核身通过率大于第二阈值的第二核

10 身产品组合。

12、如权利要求 11 所述的装置,其特征在于,所述第一核身安全处理模块包括:

第五核身通过率获取模块,被配置为将用户的属性特征输入至训练好的核身安全模型,得到与该用户的属性特征相对应的第一核身产品的第五核身通过率;

15 第六核身通过率获取模块,被配置为统计目标时间范围内的与该用户的属性特征相对应的第一核身产品的第六核身通过率;

第一核身通过率确定模块,被配置为将第一核身产品的第五核身通过率和第六核身通过率进行比较,取较大值作为第一核身产品的第一核身通过率。

20 13、如权利要求 12 所述的装置,其特征在于,所述第六核身通过率获取模块被配置为将目标时间范围内的风险分为风险群组,统计每个风险群组的与所述第一核身产品对应的核身通过率,将各个风险群组的所述核身通过率加权,得到第一核身产品的第六核身通过率。

14、如权利要求 11 所述的装置,其特征在于,还包括:

筛选模块,被配置为根据预先设置的安全管控策略,对得到的第二核身产品进行筛选,并将筛选后的第二核身产品输入至训练好的核身适用模型。

25 15、如权利要求 11 所述的装置,其特征在于,所述第二核身安全处理模块通过以下公式计算所述第一核身产品组合的第二核身通过率:

$$P(A+B)=1-(1-P(A))*(1-P(B))$$

其中, $P(A+B)$ 为第一核身产品组合的第二核身通过率;

$P(A)$ 、 $P(B)$ 分别为第一核身产品的第一核身通过率。

30 16、如权利要求 11 所述的装置,其特征在于,所述第二核身适用处理模块通过以下公式计算所述第二核身产品组合的第四核身通过率:

$$P(C+D)=1-(1-P(C))*(1-P(D))$$

其中， $P(C+D)$ 为第二核身产品组合的第四核身通过率；

$P(C)$ 、 $P(D)$ 均为第二核身产品的第三核身通过率。

17、一种核身安全模型的生成装置，其特征在于，包括：

- 5 第一获取模块，被配置为获取核身安全样本和第一标签，所述核身安全样本包括历史突破事件、该历史突破事件对应的核身产品和用户的历史属性特征，所述第一标签为所述历史突破事件对应的核身产品核身通过或不通过；

 第一训练模块，被配置为建立核身安全模型，将核身安全样本和第一标签输入至所述核身安全模型进行训练；

- 10 第一存储模块，被配置为保存所述核身安全模型。

 18、如权利要求 17 所述的装置，其特征在于，所述核身安全模型为马尔可夫链蒙特卡罗 MCMC 模型。

 19、一种核身适用模型的生成装置，其特征在于，包括：

- 第二获取模块，被配置为获取核身适用样本和第二标签，所述核身适用样本包括历史非突破事件、该历史非突破事件对应的核身产品和用户的历史属性特征，所述第二标签为所述历史非突破事件对应的核身产品核身通过或不通过；
- 15 第二训练模块，被配置为建立核身适用模型，将核身适用样本和第二标签输入至所述核身适用模型进行训练；

 第二存储模块，被配置为保存所述核身适用模型。

- 20 20、如权利要求 19 所述的装置，其特征在于，所述核身适用模型为随机森林算法模型。

 21、一种计算设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机指令，其特征在于，所述处理器执行所述指令时实现该指令被处理器执行时实现权利要求 1-6 或者 7-8 或者 9-10 任意一项所述方法的步骤。

- 25 22、一种计算机可读存储介质，其存储有计算机指令，其特征在于，该指令被处理器执行时实现权利要求 1-6 或者 7-8 或者 9-10 任意一项所述方法的步骤。