



# (12)发明专利申请

(10)申请公布号 CN 108182000 A

(43)申请公布日 2018.06.19

(21)申请号 201711418534.6

(22)申请日 2017.12.25

(71)申请人 科大讯飞股份有限公司

地址 230000 安徽省合肥市高新开发区望江西路666号

(72)发明人 王夏鸣 翟吉博

(74)专利代理机构 深圳市华优知识产权代理事务所(普通合伙) 44319

代理人 余薇

(51)Int.Cl.

G06F 3/023(2006.01)

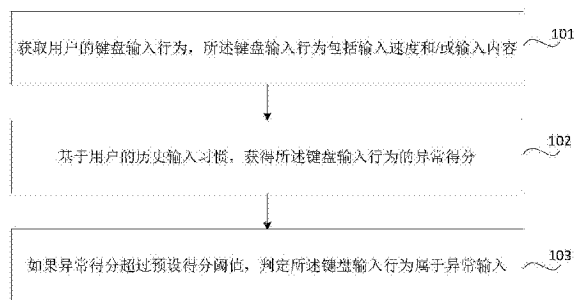
权利要求书3页 说明书8页 附图3页

## (54)发明名称

键盘输入检测方法及装置、存储介质、电子设备

## (57)摘要

本公开提供一种键盘输入检测方法及装置、存储介质、电子设备。该方法包括：获取用户的键盘输入行为，所述键盘输入行为包括输入速度和/或输入内容；基于用户的历史输入习惯，获得所述键盘输入行为的异常得分；如果所述异常得分超过预设得分阈值，判定所述键盘输入行为属于异常输入。如此方案，可以实现异常键盘输入检测，且不需要对输入内容、字符长度、所用语种等做特殊限制。



1. 一种键盘输入检测方法,其特征在于,所述方法包括:

获取用户的键盘输入行为,所述键盘输入行为包括输入速度和/或输入内容;

基于用户的历史输入习惯,获得所述键盘输入行为的异常得分;

如果所述异常得分超过预设得分阈值,判定所述键盘输入行为属于异常输入。

2. 根据权利要求1所述的方法,其特征在于,所述输入速度包括:单个字符的键入速度 $v$ 、基于所述键入速度 $v$ 进行字符输入的持续时长 $t$ 、以及基于 $(v, t)$ 进行字符输入的行为在指定时间段 $T$ 内出现的频次 $n$ ;

所述基于用户的历史输入习惯,获得所述键盘输入行为的异常得分,包括:

判断所述频次 $n$ 是否超过基于所述历史输入习惯确定出的预设频次 $n_0$ ;

如果所述频次 $n$ 超过所述预设频次 $n_0$ ,则计算所述指定时间段 $T$ 内正常出现频次 $n$ 的概率 $P_{nT}$ ;

基于所述概率 $P_{nT}$ ,获得所述键盘输入行为的异常得分 $S=1-P_{nT}$ 。

3. 根据权利要求1所述的方法,其特征在于,所述输入内容包括:在指定时间段 $T$ 内键入的字符串,所述字符串包括 $M$ 个字符;

所述基于用户的历史输入习惯,获得所述键盘输入行为的异常得分,包括:

基于所述历史输入习惯以及语法规则构建语言模型,利用所述语言模型预测所述字符串属于正常输入的得分 $Q(C_T)$ ;基于所述得分 $Q(C_T)$ ,获得所述键盘输入行为的异常得分 $S=1-Q(C_T)$ ;

或者,

提取所述字符串的音节切分特征,所述音节切分特征包括所述字符串的音节切分量 $N$ ,和/或所述字符串的音节切分比例 $D=N/M$ ;判断所述音节切分特征是否超过基于所述历史输入习惯确定的预设切分特征;如果所述音节切分特征超过所述预设切分特征,则计算所述指定时间段 $T$ 内正常得到所述音节切分特征的概率 $P_T$ ;基于所述概率 $P_T$ ,获得所述键盘输入行为的异常得分 $S=1-P_T$ 。

4. 根据权利要求1所述的方法,其特征在于,所述输入内容包括:在指定时间段 $T$ 内键入的字符串,所述字符串包括 $M$ 个字符;

所述基于用户的历史输入习惯,获得所述键盘输入行为的异常得分,包括:

基于所述历史输入习惯以及语法规则构建语言模型,利用所述语言模型预测所述字符串属于正常输入的得分 $Q(C_T)$ ;

提取所述字符串的音节切分特征,所述音节切分特征包括所述字符串的音节切分量 $N$ ,和/或所述字符串的音节切分比例 $D=N/M$ ;判断所述音节切分特征是否超过基于所述历史输入习惯确定的预设切分特征;如果所述音节切分特征超过所述预设切分特征,则计算所述指定时间段 $T$ 内正常得到所述音节切分特征的概率 $P_T$ ;

基于所述得分 $Q(C_T)$ 以及所述概率 $P_T$ ,获得所述键盘输入行为的异常得分 $S=A(1-Q(C_T))+B(1-P_T)$ , $A$ 、 $B$ 为经验参数。

5. 根据权利要求1至4任一项所述的方法,其特征在于,所述判定所述键盘输入行为属于异常输入之后,所述方法还包括:

向用户进行异常输入提醒,若用户确认所述键盘输入行为属于正常行为,则将所述键盘输入行为确定为所述历史输入习惯。

6. 一种键盘输入检测装置,其特征在于,所述装置包括:

键盘输入行为获取模块,用于获取用户的键盘输入行为,所述键盘输入行为包括输入速度和/或输入内容;

异常得分获得模块,用于基于用户的历史输入习惯,获得所述键盘输入行为的异常得分;

异常输入判定模块,用于在所述异常得分超过预设得分阈值时,判定所述键盘输入行为属于异常输入。

7. 根据权利要求6所述的装置,其特征在于,所述输入速度包括:单个字符的键入速度 $v$ 、基于所述键入速度 $v$ 进行字符输入的持续时长 $t$ 、以及基于 $(v, t)$ 进行字符输入的行为在指定时间段 $T$ 内出现的频次 $n$ ,则

所述异常得分获得模块,用于判断所述频次 $n$ 是否超过基于所述历史输入习惯确定出的预设频次 $n_0$ ;如果所述频次 $n$ 超过所述预设频次 $n_0$ ,则计算所述指定时间段 $T$ 内正常出现频次 $n$ 的概率 $P_{nr}$ ;基于所述概率 $P_{nr}$ ,获得所述键盘输入行为的异常得分 $S=1-P_{nr}$ 。

8. 根据权利要求6所述的装置,其特征在于,所述输入内容包括:在指定时间段 $T$ 内键入的字符串,所述字符串包括 $M$ 个字符;

所述异常得分获得模块,用于基于所述历史输入习惯以及语法规则构建语言模型,利用所述语言模型预测所述字符串属于正常输入的得分 $Q(C_T)$ ;基于所述得分 $Q(C_T)$ ,获得所述键盘输入行为的异常得分 $S=1-Q(C_T)$ ;

或者,

所述异常得分获得模块,用于提取所述字符串的音节切分特征,所述音节切分特征包括所述字符串的音节切分量 $N$ ,和/或所述字符串的音节切分比例 $D=N/M$ ;判断所述音节切分特征是否超过基于所述历史输入习惯确定的预设切分特征;如果所述音节切分特征超过所述预设切分特征,则计算所述指定时间段 $T$ 内正常得到所述音节切分特征的概率 $P_T$ ;基于所述概率 $P_T$ ,获得所述键盘输入行为的异常得分 $S=1-P_T$ 。

9. 根据权利要求6所述的装置,其特征在于,所述输入内容包括:在指定时间段 $T$ 内键入的字符串,所述字符串包括 $M$ 个字符,所述异常得分获得模块包括:

语音模型得分预测模块,用于基于所述历史输入习惯以及语法规则构建语言模型,利用所述语言模型预测所述字符串属于正常输入的得分 $Q(C_T)$ ;

音节切分特征概率计算模块,用于提取所述字符串的音节切分特征,所述音节切分特征包括所述字符串的音节切分量 $N$ ,和/或所述字符串的音节切分比例 $D=N/M$ ;判断所述音节切分特征是否超过基于所述历史输入习惯确定的预设切分特征;如果所述音节切分特征超过所述预设切分特征,则计算所述指定时间段 $T$ 内正常得到所述音节切分特征的概率 $P_T$ ;

异常得分获得子模块,用于基于所述得分 $Q(C_T)$ 以及所述概率 $P_T$ ,获得所述键盘输入行为的异常得分 $S=A(1-Q(C_T))+B(1-P_T)$ , $A$ 、 $B$ 为经验参数。

10. 根据权利要求6至9任一项所述的装置,其特征在于,所述装置还包括:

正常行为确认模块,用于在所述异常输入判定模块判定所述键盘输入行为属于异常输入时,向用户进行异常输入提醒,若用户确认所述键盘输入行为属于正常行为,则将所述键盘输入行为确定为所述历史输入习惯。

11. 一种存储设备,其中存储有多条指令,其特征在于,所述指令由处理器加载,执行权

利要求1至5任一项所述方法的步骤。

12. 一种电子设备,其特征在于,所述电子设备包括;  
权利要求11所述的存储设备;以及  
处理器,用于执行所述存储设备中的指令。

## 键盘输入检测方法、装置、存储介质、电子设备

### 技术领域

[0001] 本公开涉及人机交互技术领域,具体地,涉及一种键盘输入检测方法、装置、存储介质、电子设备。

### 背景技术

[0002] 目前,键盘输入作为最常用的人机交互方式,被广泛应用于各种设备。在键盘处于使用状态时,可能会因为突发情况被触发,误输入一些内容,例如,放在口袋的手机,在未锁屏的情况下误碰了键盘;放在饭桌上的手机,在未锁屏的情况下被溅上油渍水渍,误触发虚拟键盘;幼儿或宠物乱按键盘,输入一些无意义的字符串,等等。

[0003] 通常,设备只要监测到有输入,就会记录输入内容,并在误触发发送按键时,发送之前记录的内容,如此势必会降低用户的使用体验。对应于此,现有技术提供如下识别异常键盘输入的方案:

[0004] 根据键盘输入行为生成多个维度的Z值,维度指的是每个字符的敲击时间和/或两个连续字符的间隔时间,每个维度均符合正态分布;然后利用每个维度的Z值进行平方求和,得到键盘输入行为对应的卡方值;最后通过比较卡方值与预设阈值大小的方式,判断键盘输入是否异常。这种方案,只适用于输入内容固定的键盘,例如,账号、密码、邮件键盘,不适用可以输入任何内容、任意长度的普通键盘。

### 发明内容

[0005] 本公开的主要目的是提供一种键盘输入检测方法、装置、存储介质、电子设备,有助于检测异常键盘输入,且不需要对输入内容、字符长度、所用语种等做特殊要求。

[0006] 为了实现上述目的,本公开提供一种键盘输入检测方法,所述方法包括:

[0007] 获取用户的键盘输入行为,所述键盘输入行为包括输入速度和/或输入内容;

[0008] 基于用户的历史输入习惯,获得所述键盘输入行为的异常得分;

[0009] 如果所述异常得分超过预设得分阈值,判定所述键盘输入行为属于异常输入。

[0010] 可选地,所述输入速度包括:单个字符的键入速度 $v$ 、基于所述键入速度 $v$ 进行字符输入的持续时长 $t$ 、以及基于 $(v, t)$ 进行字符输入的行为在指定时间段 $T$ 内出现的频次 $n$ ;

[0011] 所述基于用户的历史输入习惯,获得所述键盘输入行为的异常得分,包括:

[0012] 判断所述频次 $n$ 是否超过基于所述历史输入习惯确定出的预设频次 $n_0$ ;

[0013] 如果所述频次 $n$ 超过所述预设频次 $n_0$ ,则计算所述指定时间段 $T$ 内正常出现频次 $n$ 的概率 $P_{nT}$ ;

[0014] 基于所述概率 $P_{nT}$ ,获得所述键盘输入行为的异常得分 $S=1-P_{nT}$ 。

[0015] 可选地,所述输入内容包括:在指定时间段 $T$ 内键入的字符串,所述字符串包括 $M$ 个字符;

[0016] 所述基于用户的历史输入习惯,获得所述键盘输入行为的异常得分,包括:

[0017] 基于所述历史输入习惯以及语法规则构建语言模型,利用所述语言模型预测所述

字符串属于正常输入的得分 $Q(C_T)$ ;基于所述得分 $Q(C_T)$ ,获得所述键盘输入行为的异常得分 $S=1-Q(C_T)$ ;

[0018] 或者,

[0019] 提取所述字符串的音节切分特征,所述音节切分特征包括所述字符串的音节切分量 $N$ ,和/或所述字符串的音节切分比例 $D=N/M$ ;判断所述音节切分特征是否超过基于所述历史输入习惯确定的预设切分特征;如果所述音节切分特征超过所述预设切分特征,则计算所述指定时间段 $T$ 内正常得到所述音节切分特征的概率 $P_T$ ;基于所述概率 $P_T$ ,获得所述键盘输入行为的异常得分 $S=1-P_T$ 。

[0020] 可选地,所述输入内容包括:在指定时间段 $T$ 内键入的字符串,所述字符串包括 $M$ 个字符;

[0021] 所述基于用户的历史输入习惯,获得所述键盘输入行为的异常得分,包括:

[0022] 基于所述历史输入习惯以及语法规则构建语言模型,利用所述语言模型预测所述字符串属于正常输入的得分 $Q(C_T)$ ;

[0023] 提取所述字符串的音节切分特征,所述音节切分特征包括所述字符串的音节切分量 $N$ ,和/或所述字符串的音节切分比例 $D=N/M$ ;判断所述音节切分特征是否超过基于所述历史输入习惯确定的预设切分特征;如果所述音节切分特征超过所述预设切分特征,则计算所述指定时间段 $T$ 内正常得到所述音节切分特征的概率 $P_T$ ;

[0024] 基于所述得分 $Q(C_T)$ 以及所述概率 $P_T$ ,获得所述键盘输入行为的异常得分 $S=A(1-Q(C_T))+B(1-P_T)$ , $A$ 、 $B$ 为经验参数。

[0025] 可选地,所述判定所述键盘输入行为属于异常输入之后,所述方法还包括:

[0026] 向用户进行异常输入提醒,若用户确认所述键盘输入行为属于正常行为,则将所述键盘输入行为确定为所述历史输入习惯。

[0027] 本公开提供一种键盘输入检测装置,所述装置包括:

[0028] 键盘输入行为获取模块,用于获取用户的键盘输入行为,所述键盘输入行为包括输入速度和/或输入内容;

[0029] 异常得分获得模块,用于基于用户的历史输入习惯,获得所述键盘输入行为的异常得分;

[0030] 异常输入判定模块,用于在所述异常得分超过预设得分阈值时,判定所述键盘输入行为属于异常输入。

[0031] 可选地,所述输入速度包括:单个字符的键入速度 $v$ 、基于所述键入速度 $v$ 进行字符输入的持续时长 $t$ 、以及基于 $(v, t)$ 进行字符输入的行为在指定时间段 $T$ 内出现的频次 $n$ ,则

[0032] 所述异常得分获得模块,用于判断所述频次 $n$ 是否超过基于所述历史输入习惯确定出的预设频次 $n_0$ ;如果所述频次 $n$ 超过所述预设频次 $n_0$ ,则计算所述指定时间段 $T$ 内正常出现频次 $n$ 的概率 $P_{nT}$ ;基于所述概率 $P_{nT}$ ,获得所述键盘输入行为的异常得分 $S=1-P_{nT}$ 。

[0033] 可选地,所述输入内容包括:在指定时间段 $T$ 内键入的字符串,所述字符串包括 $M$ 个字符;

[0034] 所述异常得分获得模块,用于基于所述历史输入习惯以及语法规则构建语言模型,利用所述语言模型预测所述字符串属于正常输入的得分 $Q(C_T)$ ;基于所述得分 $Q(C_T)$ ,获得所述键盘输入行为的异常得分 $S=1-Q(C_T)$ ;

[0035] 或者,

[0036] 所述异常得分获得模块,用于提取所述字符串的音节切分特征,所述音节切分特征包括所述字符串的音节切分量 $N$ ,和/或所述字符串的音节切分比例 $D=N/M$ ;判断所述音节切分特征是否超过基于所述历史输入习惯确定的预设切分特征;如果所述音节切分特征超过所述预设切分特征,则计算所述指定时间段 $T$ 内正常得到所述音节切分特征的概率 $P_T$ ;基于所述概率 $P_T$ ,获得所述键盘输入行为的异常得分 $S=1-P_T$ 。

[0037] 可选地,所述输入内容包括:在指定时间段 $T$ 内键入的字符串,所述字符串包括 $M$ 个字符,所述异常得分获得模块包括:

[0038] 语音模型得分预测模块,用于基于所述历史输入习惯以及语法规则构建语言模型,利用所述语言模型预测所述字符串属于正常输入的得分 $Q(C_T)$ ;

[0039] 音节切分特征概率计算模块,用于提取所述字符串的音节切分特征,所述音节切分特征包括所述字符串的音节切分量 $N$ ,和/或所述字符串的音节切分比例 $D=N/M$ ;判断所述音节切分特征是否超过基于所述历史输入习惯确定的预设切分特征;如果所述音节切分特征超过所述预设切分特征,则计算所述指定时间段 $T$ 内正常得到所述音节切分特征的概率 $P_T$ ;

[0040] 异常得分获得子模块,用于基于所述得分 $Q(C_T)$ 以及所述概率 $P_T$ ,获得所述键盘输入行为的异常得分 $S=A(1-Q(C_T))+B(1-P_T)$ , $A$ 、 $B$ 为经验参数。

[0041] 可选地,所述装置还包括:

[0042] 正常行为确认模块,用于在所述异常输入判定模块判定所述键盘输入行为属于异常输入时,向用户进行异常输入提醒,若用户确认所述键盘输入行为属于正常行为,则将所述键盘输入行为确定为所述历史输入习惯。

[0043] 本公开提供一种存储设备,其中存储有多条指令,所述指令由处理器加载,执行上述键盘输入检测方法的步骤。

[0044] 本公开提供一种电子设备,所述电子设备包括:

[0045] 上述的存储设备;以及

[0046] 处理器,用于执行所述存储设备中的指令。

[0047] 本公开方案中,可以在监测到用户通过键盘进行人机交互时,获取用户的键盘输入行为,并基于用户的历史输入习惯,计算键盘输入行为的异常得分,实现异常输入检测。具体地,如果异常得分超过预设得分阈值,则说明键盘输入行为不符合用户的历史输入习惯,可以判定该键盘输入行为属于异常输入。此外,结合上文介绍可知,本公开方案不需要对输入内容、字符长度、所用语种等做特殊限制,且适用于各种类型的键盘。

[0048] 本公开的其他特征和优点将在随后的具体实施方式部分予以详细说明。

## 附图说明

[0049] 附图是用来提供对本公开的进一步理解,并且构成说明书的一部分,与下面的具体实施方式一起用于解释本公开,但并不构成对本公开的限制。在附图中:

[0050] 图1为本公开方案键盘输入检测方法的流程示意图;

[0051] 图2为本公开方案中获得键盘输入行为的异常得分的流程示意图;

[0052] 图3为本公开方案中频次的正态分布示意图;

[0053] 图4为本公开方案键盘输入检测装置的构成示意图；

[0054] 图5为本公开方案用于键盘输入检测的电子设备的结构示意图。

### 具体实施方式

[0055] 以下结合附图对本公开的具体实施方式进行详细说明。应当理解的是，此处所描述的具体实施方式仅用于说明和解释本公开，并不用于限制本公开。

[0056] 参见图1，示出了本公开键盘输入检测方法的流程示意图。可以包括以下步骤：

[0057] S101，获取用户的键盘输入行为，所述键盘输入行为包括输入速度和/或输入内容。

[0058] 通常，用户的输入行为具有一定的规律性，即，用户正常的键盘输入行为应与用户以往的输入习惯相符，故本公开方案监测到用户通过键盘进行内容输入时，可以获取该用户的键盘输入行为，并基于用户的历史输入习惯，实现异常输入检测。

[0059] 作为一种示例，本公开方案中的键盘输入行为可以体现为输入速度和/或输入内容，具体可参见下文键盘输入行为的异常得分处所做介绍，此处暂不详述。

[0060] S102，基于用户的历史输入习惯，获得所述键盘输入行为的异常得分。

[0061] 结合键盘输入行为的不同表现形式，本公开方案的异常得分可以体现为以下三种情况中的任一种，下面分别进行解释说明。

[0062] 1. 键盘输入行为可以是输入速度

[0063] 作为一种示例，输入速度可以包括：单个字符的键入速度 $v$ 、基于键入速度 $v$ 进行字符输入的持续时长 $t$ 、以及指定时间段 $T$ 内基于 $(v, t)$ 进行字符输入的行为出现的频次 $n$ 。

[0064] 举例来说，可以通过输入法记录每个字符的键入时间，利用相邻两个键入时间之间的间隔，计算单个字符的键入速度 $v$ ，即键入速度 $v$ 属于瞬时速度。以键入“xiguan”为例，可以记录“x”、“i”的键入时间 $t_x$ 、 $t_i$ ，利用 $(t_i - t_x)$ 计算字符“x”的键入速度 $v_x$ ，例如 $(t_i - t_x) = 0.5s$ ，则字符“x”的键入速度 $v_x = 2$ 键/s；同理，可以利用 $(t_g - t_i)$ 计算字符“i”的键入速度 $v_i$ ，以此类推，可以获得每个字符对应的键入速度 $v_g$ 、 $v_u$ 、 $v_a$ 、 $v_n$ 。可以理解地，最后一个字符“n”的键入速度 $v_n$ 可以忽略不计；或者；可以设定为默认速度 $v_0$ ；或者，可以与前一字符的输入速度相同，即 $v_n = v_a$ ，本公开方案对此可不做具体限定，可以结合实际应用需求确定。

[0065] 仍以键入“xiguan”为例，如果 $v_x = v_i = v_g = v_1$ 、 $v_u = v_2$ ，则可计算得到基于键入速度 $v_1$ 进行字符输入的持续时长 $t_1$ 。可以理解地，如果用户连续键入字符串“xiguan”，则 $t_1 = (t_g - t_x)$ ；如果用户在键入字符串的时候出现了中断，例如，键入“xi”后进行上屏候选找到对应的汉字“习”、点击发送按键等，可以记录中断的持续时长 $t_{\text{中断}}$ ，对应地 $t_1 = (t_g - t_x - t_{\text{中断}})$ 。即， $t_1$ 仅指基于键入速度 $v_1$ 进行字符输入的持续时长。

[0066] 按照上文所做介绍，可以统计得到基于 $(v, t)$ 进行字符输入的行为出现的频次 $n$ 。举例来说，如果指定时间段 $T = 30s$ ， $(2, 2.5, 10)$ 可以理解地为，键入速度为2键/s且按照该键入速度进行字符输入的持续时长为2.5s，这样的情况在30s内出现了10次。

[0067] 作为一种示例，可以根据实际使用需求定义指定时间段的时长，本公开方案对此可不做具体限定。通常，指定时间段不超过键盘输入行为对应的采集时长。

[0068] 获得输入速度的相关信息后，可以按照图2所示流程图得到键盘输入行为的异常得分，具体可以包括以下步骤：

[0069] S201,判断所述频次 $n$ 是否超过基于所述历史输入习惯确定出的预设频次 $n_0$ 。

[0070] 作为一种示例,可以按照以下方法确定预设频次 $n_0$ :

[0071] 可以绘制总字符样本为 $H$ 的情况下,所有(键入速度,持续时长)对应频次的概率分布,以此来表示用户的历史输入习惯,通常,用户的历史输入习惯是一定的,即历史输入习惯符合正态分布。举例来说,图3示出了 $(v_1, t_1)$ 组合下 $n$ 的正态分布示意图,其中,横轴表示频次 $n$ ,纵轴表示正常出现 $n$ 的概率。

[0072] 举例来说,可以设定正态分布80%面积对应的是预设频次 $n_0$ ,可以将频次最低的10%视为输入过慢、将频次最高的10%视为输入过快,两者均可视为异常输入。具体可以利用如下公式求出 $n_0$ :

$$[0073] \int_{n-2\mu}^n \frac{1}{\sqrt{2\pi\sigma}} e^{\left(-\frac{(n-\mu)^2}{2\sigma^2}\right)} dn = 80\%$$

[0074] 对于预设频次对应的面积、输入过慢对应的频次、输入过快对应的频次、异常输入的具体表现形式,本公开方案可不做限定,具体可结合实际应用需求确定。

[0075] S202,如果所述频次 $n$ 与所述预设频次 $n_0$ 不符,则计算所述指定时间段 $T$ 内正常出现频次 $n$ 的概率 $P_{nT}$ 。

[0076] S203,基于所述概率 $P_{nT}$ ,获得所述键盘输入行为的异常得分 $S=1-P_{nT}$ 。

[0077] 作为一种示例,本公开方案中 $n$ 与 $n_0$ 不符,可以理解为 $n$ 超过了输入过快对应的频次,或者, $n$ 低于输入过慢对应的频次。以图3所举示例为例,如果用户在输入过程中,出现 $(v, t)$ 对应的 $n > n_0$ 的情况,则认为用户进行了可疑的异常快速输入,可以按照以下公式计算指定时间段 $T$ 内正常出现频次 $n$ 的概率 $P_{nT}$ ,进而得到输入速度异常的得分 $S=1-P_{nT}$ ,可以作为键盘输入行为的异常得分。

$$[0078] P_{nT} = \frac{1}{\sqrt{2\pi\sigma}} e^{\left(-\frac{(n-\mu)^2}{2\sigma^2}\right)}$$

[0079] 2. 键盘输入行为可以是输入内容

[0080] 作为一种示例,输入内容可以包括:在指定时间段 $T$ 内键入的字符串,该字符串可以包括 $M$ 个字符, $M \geq 1$ 。

[0081] 本公开方案提供以下基于输入内容进行异常输入检测的方案:

[0082] (1) 根据字符串的语言模型得分进行异常输入检测

[0083] 作为一种示例,可以先基于历史输入习惯以及语法规则构建语言模型,如此,用户在时间段 $T$ 内输入的字符串通过输入法整句解码后的句子 $C$ ,经语言模型后,可以得到该字符串属于正常输入的得分 $Q(C_T)$ ,进而得到输入内容异常的得分 $S=1-Q(C_T)$ ,可以作为键盘输入行为的异常得分。

[0084] 通常,得分越低,意味着用户之前输入该字符串的概率越低,且从正常语法规则来看,该字符串也少有先例,输入内容异常的可能性越高。

[0085] (2) 根据字符串的音节切分特征得分进行异常输入检测

[0086] 作为一种示例,可以通过拼音引擎,对用户在时间段 $T$ 内输入的字符串进行音节切分,提取出字符串的音节切分特征,例如,音节切分特征可以包括字符串的音节切分量 $N$ ,和/或字符串的音节切分比例 $D=N/M$ ;判断音节切分特征是否超过基于历史输入习惯确定

的预设切分特征；如果音节切分特征超过预设切分特征，则可计算正常得到音节切分特征的概率 $P_T$ ，进而得到输入内容异常的得分 $S=1-P_T$ ，可以作为键盘输入行为的异常得分。

[0087] 通常，用户正常输入时，输入内容具有一定的规律性。举例来说，用户习惯短拼，如单次输入可能仅涉及一两个音节；或者，用户习惯简拼，如仅输入声母或声母的首字母，此时音节切分比例接近1:1；或者，用户输入具有个性化特点，如对词组进行部分简拼，以“上班”为例，用户习惯的输入方式为“shangb”。

[0088] 以音节切分比例为例，如果键盘胡乱输入“dhusahfa jbahda”，可以切分出如下音节：d' hu' sa' h' fa' j' ba' h' da，即14个字符切分出9个音节，通常不符合用户输入习惯，可识别为异常输入。

[0089] 对应音节切分比例，可以按照以下方法确定预设切分特征：

[0090] 可以对用户单次输入的音节切分比例进行数据积累学习，针对不同音节切分比例，建立对应的概率分布函数 $F(D)$ 。作为一种示例，可以利用80%概率对应的音节切分比例，确定预设切分特征 $D_0$ ，具体可通过如下公式求出 $D_0$ ：

$$[0091] \int_0^{D_0} F(D) = 80\%$$

[0092] 如果用户在输入过程中，出现 $D>D_0$ 的情况，则可认为用户进行了可疑的异常输入，可以基于概率分布函数 $F(D)$ ，确定出指定时间段 $T$ 内正常切分得到 $D$ 的概率 $P_T$ ，进而得到输入内容异常的得分 $S=1-P_T$ ，可以作为键盘输入行为的异常得分。

[0093] 以音节切分量为例，可以统计用户单次输入的音节切分量，针对不同音节切分量，建立对应的概率分布函数 $F(M)$ 。与音节切分比例类似，可以确定出预设切分特征 $M_0$ ，并在 $M>M_0$ 时，计算出输入内容异常的得分 $S=1-P_T$ ，作为键盘输入行为的异常得分。

[0094] 可以理解地，计算 $D_0$ 、 $M_0$ 时选取的概率可结合实际应用需求确定，本公开方案对此可不做具体限定。

[0095] (3) 根据字符串的语言模型得分和音节切分特征得分进行异常输入检测

[0096] 按照上文所做介绍，获得后 $Q(C_T)$ 、 $P_T$ ，可以基于二者得到键盘输入行为的异常得分 $S=A(1-Q(C_T))+B(1-P_T)$ ， $A$ 、 $B$ 为经验参数。具体过程可参照上文所做介绍，此处不再赘述。

[0097] 3. 键盘输入行为可以是输入速度以及输入内容

[0098] 按照上文所做介绍，获得输入速度异常的得分 $(1-P_{nT})$ 、以及输入内容异常的得分，例如 $[A(1-Q(C_T))+B(1-P_T)]$ ，可以基于二者得到键盘输入行为的异常得分 $S=a(1-P_{nT})+b[A(1-Q(C_T))+B(1-P_T)]$ ， $a$ 、 $b$ 为经验参数。具体过程可参照上文所做介绍，此处不再赘述。

[0099] S103，如果所述异常得分超过预设得分阈值，判定所述键盘输入行为属于异常输入。

[0100] 结合上文异常得分的三种情况，预设得分阈值可以体现为：

[0101] 1. 键盘输入行为的异常得分 $S=1-P_{nT}$

[0102] 可以结合用户以往键盘输入行为的输入速度，确定预设得分阈值 $S_1$ ，如果 $S>S_1$ ，则可判定此次键盘输入行为属于异常输入。

[0103] 2. 键盘输入行为的异常得分 $S=A(1-Q(C_T))+B(1-P_T)$

[0104] 可以结合用户以往键盘输入行为的输入内容，确定预设得分阈值 $S_2$ ，如果 $S>S_2$ ，则可判定此次键盘输入行为属于异常输入。

[0105] 3. 键盘输入行为的异常得分 $S=a(1-P_{nT})+b[A(1-Q(C_T))+B(1-P_T)]$

[0106] 可以结合用户以往键盘输入行为的输入速度和输入内容,确定预设得分阈值 $S_3$ ,如果 $S>S_3$ ,则可判定此次键盘输入行为属于异常输入。

[0107] 综上,本公开方案可以获取用户的键盘输入行为,并计算该键盘输入行为的异常得分,通常,异常得分越低,说明该键盘输入行为越符合用户的历史输入习惯,该键盘输入行为属于正常输入的可能性越大;反之则属于异常输入的可能性越大。此外,结合上文介绍可知,本公开方案不需要对输入内容、字符长度、所用语种等做特殊限制,且适用于各种类型的键盘,例如,可以是专用于账户、密码输入的特殊键盘,也可以是日常使用的普通键盘;或者,可以是实体键盘,也可以是虚拟键盘。

[0108] 可以理解地,本公开方案中涉及的相关阈值,可以是基于当前用户的历史键盘输入行为确定出的个性化阈值;或者,也可以是基于大多数用户的历史键盘输入行为确定出的通用阈值,本公开方案对此可不做具体限定。

[0109] 作为一种示例,本公开方案还提供一种自适应学习的优化方案,具体地,在判定用户的键盘输入行为属于异常输入之后,可以向用户进行异常输入提醒,若用户确认此次键盘输入行为属于正常行为,则可将该键盘输入行为确定为用户的历史输入习惯,用于优化输入速度和/或输入内容等相关的阈值。通过这种迭代优化方案,可以提高本公开方案异常输入检测的准确度,降低误检率。此外,若用户确认此次键盘输入行为属于异常行为,作为一种示例,可以一次性清除此次键盘输入行为输入的字符串。

[0110] 作为一种示例,在判定用户的键盘输入行为属于异常输入之后,可以继续接收用户后续键入的字符;或者,可以屏蔽用户的后续键入,并根据用户选择进行相应处理。关于键盘输入行为被判定为异常输入后的处理过程,本公开方案可不做具体限定。

[0111] 参见图4,示出了本公开键盘输入检测装置的构成示意图。所述装置可以包括:

[0112] 键盘输入行为获取模块301,用于获取用户的键盘输入行为,所述键盘输入行为包括输入速度和/或输入内容;

[0113] 异常得分获得模块302,用于基于用户的历史输入习惯,获得所述键盘输入行为的异常得分;

[0114] 异常输入判定模块303,用于在所述异常得分超过预设得分阈值时,判定所述键盘输入行为属于异常输入。

[0115] 可选地,所述输入速度包括:单个字符的键入速度 $v$ 、基于所述键入速度 $v$ 进行字符输入的持续时长 $t$ 、以及基于 $(v, t)$ 进行字符输入的行为在指定时间段 $T$ 内出现的频次 $n$ ,则

[0116] 所述异常得分获得模块,用于判断所述频次 $n$ 是否超过基于所述历史输入习惯确定出的预设频次 $n_0$ ;如果所述频次 $n$ 超过所述预设频次 $n_0$ ,则计算所述指定时间段 $T$ 内正常出现频次 $n$ 的概率 $P_{nT}$ ;基于所述概率 $P_{nT}$ ,获得所述键盘输入行为的异常得分 $S=1-P_{nT}$ 。

[0117] 可选地,所述输入内容包括:在指定时间段 $T$ 内键入的字符串,所述字符串包括 $M$ 个字符;

[0118] 所述异常得分获得模块,用于基于所述历史输入习惯以及语法规则构建语言模型,利用所述语言模型预测所述字符串属于正常输入的得分 $Q(C_T)$ ;基于所述得分 $Q(C_T)$ ,获得所述键盘输入行为的异常得分 $S=1-Q(C_T)$ ;

[0119] 或者,

[0120] 所述异常得分获得模块,用于提取所述字符串的音节切分特征,所述音节切分特征包括所述字符串的音节切分量 $N$ ,和/或所述字符串的音节切分比例 $D=N/M$ ;判断所述音节切分特征是否超过基于所述历史输入习惯确定的预设切分特征;如果所述音节切分特征超过所述预设切分特征,则计算所述指定时间段 $T$ 内正常得到所述音节切分特征的概率 $P_T$ ;基于所述概率 $P_T$ ,获得所述键盘输入行为的异常得分 $S=1-P_T$ 。

[0121] 可选地,所述输入内容包括:在指定时间段 $T$ 内键入的字符串,所述字符串包括 $M$ 个字符,所述异常得分获得模块包括:

[0122] 语音模型得分预测模块,用于基于所述历史输入习惯以及语法规则构建语言模型,利用所述语言模型预测所述字符串属于正常输入的得分 $Q(C_T)$ ;

[0123] 音节切分特征概率计算模块,用于提取所述字符串的音节切分特征,所述音节切分特征包括所述字符串的音节切分量 $N$ ,和/或所述字符串的音节切分比例 $D=N/M$ ;判断所述音节切分特征是否超过基于所述历史输入习惯确定的预设切分特征;如果所述音节切分特征超过所述预设切分特征,则计算所述指定时间段 $T$ 内正常得到所述音节切分特征的概率 $P_T$ ;

[0124] 异常得分获得子模块,用于基于所述得分 $Q(C_T)$ 以及所述概率 $P_T$ ,获得所述键盘输入行为的异常得分 $S=A(1-Q(C_T))+B(1-P_T)$ , $A$ 、 $B$ 为经验参数。

[0125] 可选地,所述装置还包括:

[0126] 正常行为确认模块,用于在所述异常输入判定模块判定所述键盘输入行为属于异常输入时,向用户进行异常输入提醒,若用户确认所述键盘输入行为属于正常行为,则将所述键盘输入行为确定为所述历史输入习惯。

[0127] 关于上述实施例中的装置,其中各个模块执行操作的具体方式已经在有关该方法的实施例中进行了详细描述,此处将不做详细阐述说明。

[0128] 参见图5,示出了本公开用于键盘输入检测的电子设备400的结构示意图。参照图5,电子设备400包括处理组件401,其进一步包括一个或多个处理器,以及由存储介质402所代表的存储设备资源,用于存储可由处理组件401的执行的指令,例如应用程序。存储介质402中存储的应用程序可以包括一个或一个以上的每一个对应于一组指令的模块。此外,处理组件401被配置为执行指令,以执行上述键盘输入检测方法。

[0129] 电子设备400还可以包括一个电源组件403,被配置为执行电子设备400的电源管理;一个有线或无线网络接口404,被配置为将电子设备400连接到网络;和一个输入输出(I/O)接口405。电子设备400可以操作基于存储在存储介质402的操作系统,例如Windows Server™,Mac OS X™,Unix™,Linux™,FreeBSD™或类似。

[0130] 以上结合附图详细描述了本公开的优选实施方式,但是,本公开并不限于上述实施方式中的具体细节,在本公开的技术构思范围内,可以对本公开的技术方案进行多种简单变型,这些简单变型均属于本公开的保护范围。

[0131] 另外需要说明的是,在上述具体实施方式中所描述的各个具体技术特征,在不矛盾的情况下,可以通过任何合适的方式进行组合,为了避免不必要的重复,本公开对各种可能的组合方式不再另行说明。

[0132] 此外,本公开的各种不同的实施方式之间也可以进行任意组合,只要其不违背本公开的思想,其同样应当视为本公开所公开的内容。

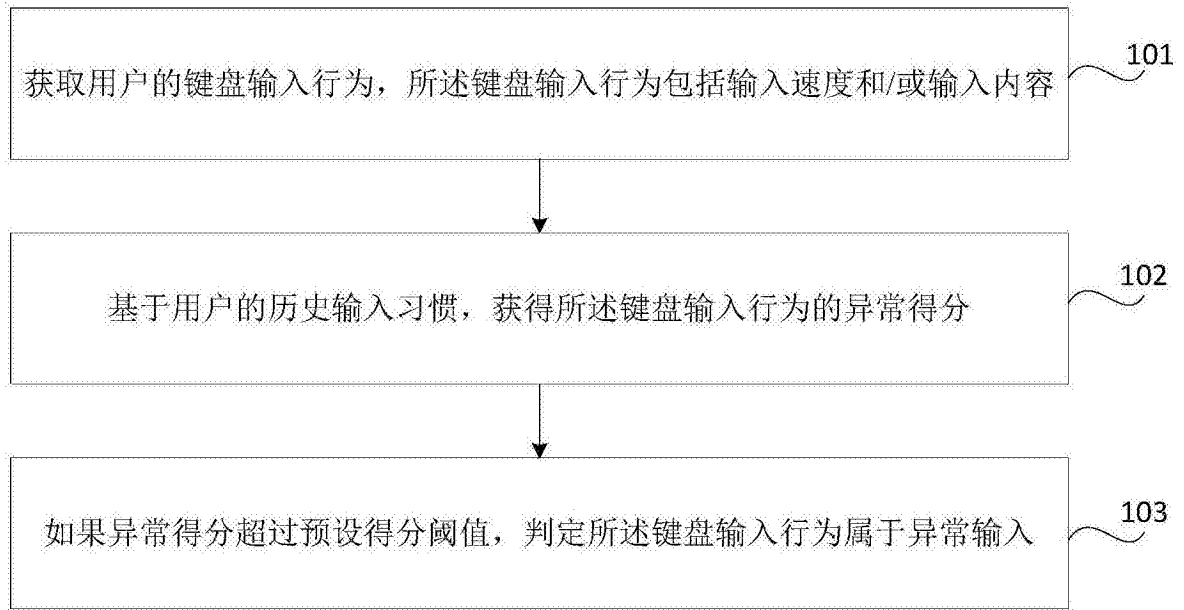


图1

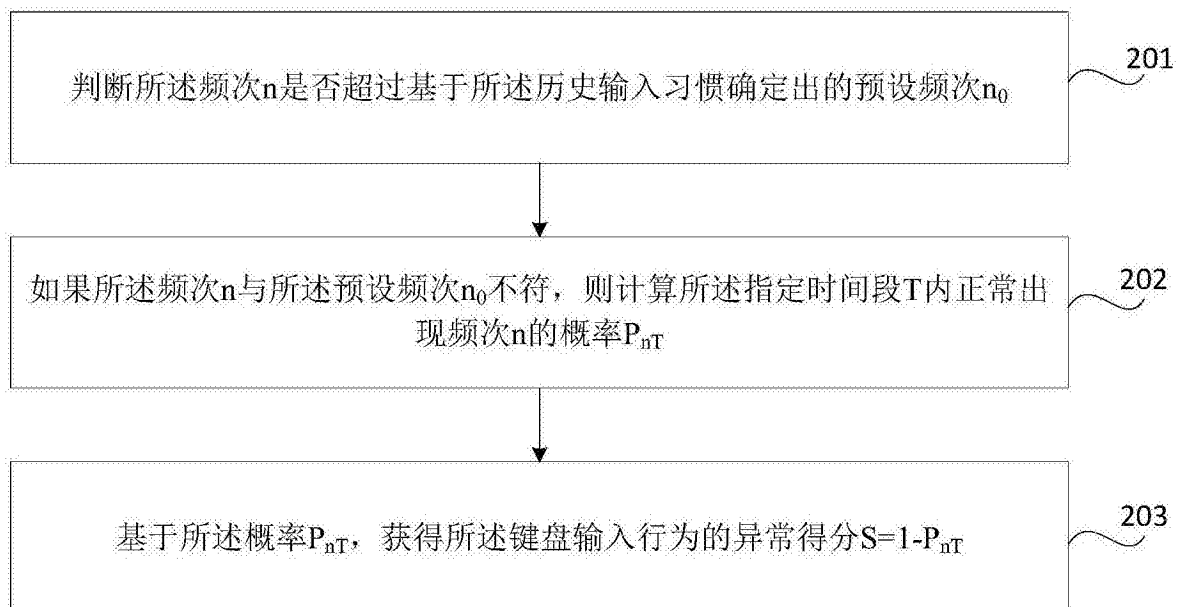


图2

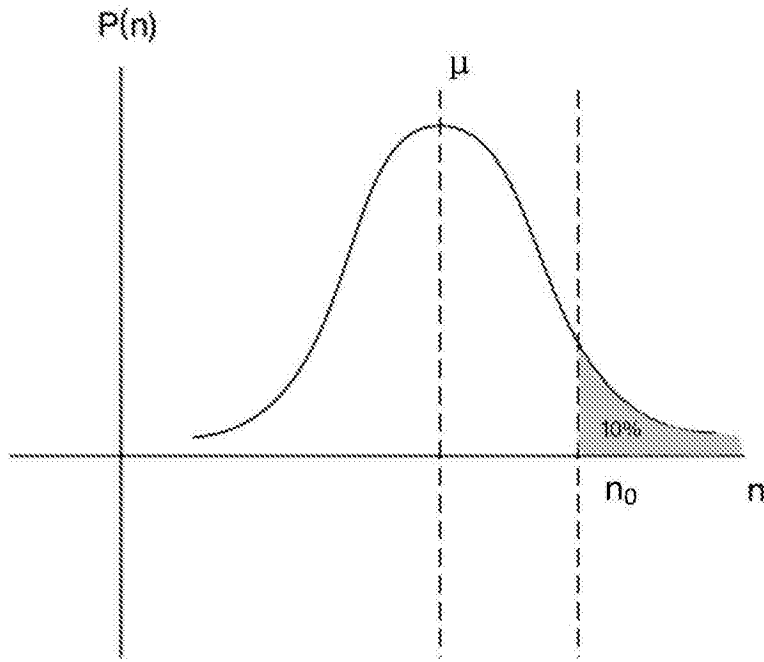


图3

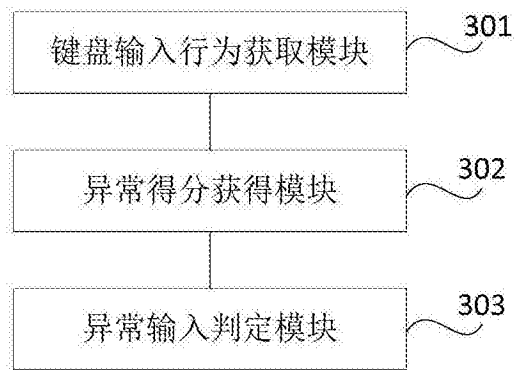


图4

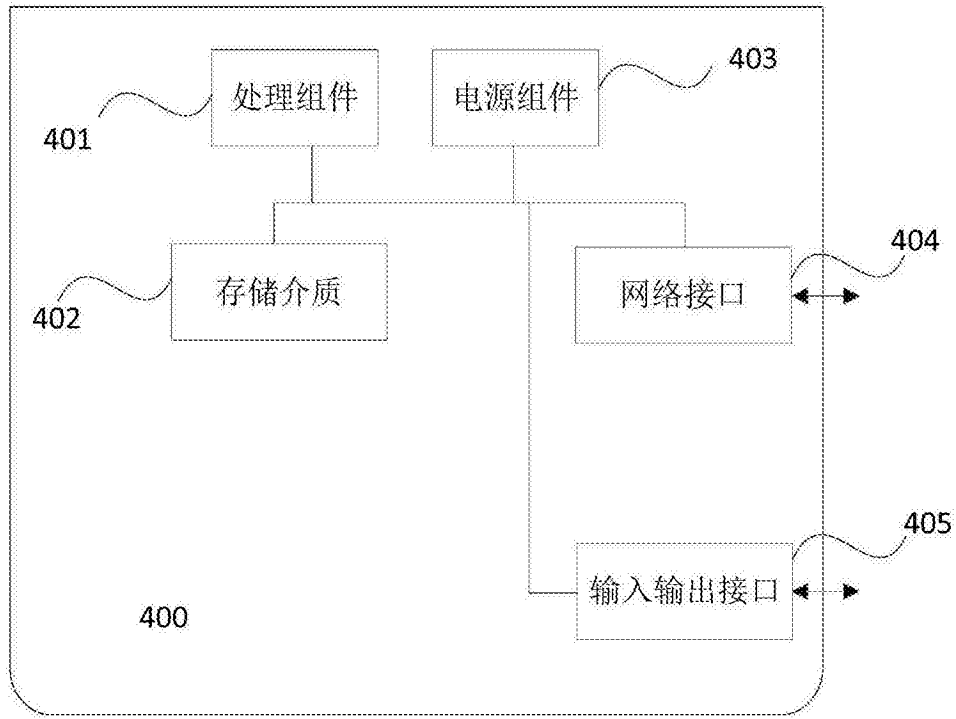


图5