



(12)发明专利

(10)授权公告号 CN 108629362 B

(45)授权公告日 2019.11.01

(21)申请号 201711445576.9

(22)申请日 2018.05.30

(65)同一申请的已公布的文献号
申请公布号 CN 108629362 A

(43)申请公布日 2018.10.09

(73)专利权人 华南师范大学
地址 510631 广东省广州市天河区中山大
道西55号

(72)发明人 李明喜 黄昌勤 文瑶 廖芳丽
韩中美

(74)专利代理机构 广州专理知识产权代理事务
所(普通合伙) 44493
代理人 谭昉

(51)Int.Cl.
G06K 9/62(2006.01)

(56)对比文件

US 2006/0280236 A1,2006.12.14,
CN 103077356 A,2013.05.01,
CN 106375369 A,2017.02.01,
CN 107766484 A,2018.03.06,

审查员 王满一

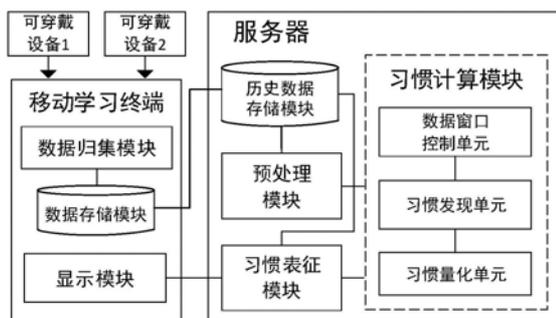
权利要求书4页 说明书12页 附图3页

(54)发明名称

一种面向移动环境的学习行为习惯发现量
化系统及方法

(57)摘要

本发明公开了一种面向移动环境的学习行
为习惯发现量化系统及方法。针对传统方法主观
性太强,时效性有限,只能测量出问卷设计者提
前设定习惯等方面的缺点,本发明拟采用基于数
据驱动的方法,实时收集移动学习环境中学习
者与平台交互的活动记录、所处位置和生理状
态等外显和内隐的学习行为习惯关联数据,并
将收集到的数据通过数据变换和数据清洗等方
法处理为便于机器识别和计算的习惯特征数
据集,进而基于聚类算法计算并发现学习行为
习惯。针对传统方法无法量化学习者学习行为
习惯强度的问题,本发明基于赫尔学习理论中
强化次数与习惯强度之间的函数关系,量化学
习行为习惯。最后,从学习行为习惯的描述、
持续时长和强度三个维度表征学习行为习惯。



1. 一种面向移动环境的学习行为习惯发现量化系统,其特征在于,系统由辅助穿戴设备,移动学习终端和服务器组成:

辅助穿戴设备,由多个可监测学习者生理状态等信息的可穿戴设备构成,负责收集学习者开展学习活动时的生理状态信息;

移动学习终端,移动学习终端宿居着一个习惯发现及量化系统的客户端,负责收集网络学习平台上的交互活动记录和学习者位置信息,归集与临时存储平台交互活动记录、学习者位置和生理状态的学习行为习惯关联数据,显示学习行为习惯的发现及量化结果,它包括如下三个模块:

数据归集模块,它连接数据存储模块,用于归集从移动学习环境中收集到的平台交互活动记录、学习者位置和生理状态的学习行为习惯关联数据;

数据存储模块,它连接数据归集模块和服务器中的历史数据存储模块,用于临时存储数据归集模块归集的学习行为习惯关联数据;

显示模块,它连接服务器中的习惯表征模块,用于显示学习行为习惯的发现及量化结果;

服务器,服务器上运行着一个习惯发现及量化系统的服务端,负责持久存储学习行为习惯关联数据和学习行为习惯表征数据,预处理与计算习惯关联数据以发现和量化学习者当前的学习行为习惯,它包括如下四个模块:

历史数据存储模块,它连接数据存储模块、预处理模块和习惯表征模块,用于接收和持久存储移动学习终端数据存储模块传来的学习行为习惯关联数据,并传送至预处理模块,此外,还存储习惯表征模块的数据;

预处理模块,它连接历史数据存储模块和习惯计算模块,用于将从历史数据存储模块中接收的学习行为习惯关联数据处理为便于机器识别和计算的习惯特征数据集;

习惯计算模块,它连接预处理模块和习惯表征模块,用于控制数据窗口的大小、计算并发现学习者的学习行为习惯簇、量化学习者当前各学习行为习惯的强度;

习惯表征模块,它连接习惯计算模块、历史数据存储模块和显示模块,用于从学习行为习惯的描述、持续时长和强度三个维度表征学习行为习惯。

2. 如权利要求1所述的一种面向移动环境的学习行为习惯发现量化系统,其特征在于,所述习惯计算模块还包括:

数据窗口控制单元,它连接习惯发现单元,用于控制数据窗口的大小,当无法准确判断某候选行为习惯簇是否为学习行为习惯簇时,将调整数据窗口的大小,以进一步判断其是否为学习行为习惯簇;

习惯发现单元,它连接数据窗口控制单元和习惯量化单元,用于计算数据窗口中的习惯特征数据,发现并获取学习者当前的学习行为习惯簇集合;

习惯量化单元,它连接习惯发现单元,用于量化学习者当前各学习行为习惯的强度。

3. 如权利要求2所述的一种面向移动环境的学习行为习惯发现量化系统,其特征在于,所述的数据窗口控制单元在习惯特征数据集X上基于时间段TW的数据窗口定义一个X的子集XW作为数据窗口的内容,TW为数据窗口的大小,TW满足以下公式:

$$TW = \begin{cases} 3T, \\ 4T, & \alpha_1 > \text{Fre}(\mathbf{HX}_{p,q}) > \alpha_2 \cap \beta_1 > \text{Det}(\mathbf{HX}_{p,q}) > \beta_2 \end{cases}$$

默认情况下基于时间的数据窗口大小 $TW=3T$,但当 $\alpha_1 < \text{Fre}(HX_{p,q}) < \alpha_2 \cap \beta_1 < \text{Den}(HX_{p,q}) < \beta_2$ 时,认为此时无法准确判断候选行为习惯簇 $HX_{p,q}$ 是否为学习行为习惯簇,于是,将时间段 TW 向前扩大一个计算周期,使 $TW=4T$,对 $HX_{p,q}$ 再次进行判断:

若计算结果为 $\text{Fre}(HX_{p,q}) \geq \alpha_2 \cap \text{Den}(HX_{p,q}) \geq \beta_2$,认为候选行为习惯簇 $HX_{p,q}$ 是学习行为习惯簇;

若计算结果为 $\text{Fre}(HX_{p,q}) \leq \alpha_1 \cap \text{Den}(HX_{p,q}) \leq \beta_1$,认为候选行为习惯簇 $HX_{p,q}$ 不是学习行为习惯簇;

若计算结果为 $\alpha_1 < \text{Fre}(HX_{p,q}) < \alpha_2 \cap \beta_1 < \text{Den}(HX_{p,q}) < \beta_2$,则将基于时间的数据窗口向后扩大一个计算周期,等下一个计算周期到达时,再对此候选行为习惯簇 $HX_{p,q}$ 进行习惯发现计算,

若计算结果为 $\text{Fre}(HX_{p,q}) \geq \alpha_2 \cap \text{Den}(HX_{p,q}) \geq \beta_2$,认为候选行为习惯簇 $HX_{p,q}$ 是学习行为习惯簇;

若计算结果为 $\text{Fre}(HX_{p,q}) \leq \alpha_1 \cap \text{Den}(HX_{p,q}) \leq \beta_1$,或者 $\alpha_1 < \text{Fre}(HX_{p,q}) < \alpha_2 \cap \beta_1 < \text{Den}(HX_{p,q}) < \beta_2$,则认为候选行为习惯簇 $HX_{p,q}$ 不是学习行为习惯簇,重新按默认情况下 $TW=3T$ 进行习惯发现计算;

其中,定义 α_1, α_2 为候选行为习惯簇发生频率的阈值,定义 β_1, β_2 为候选行为习惯簇密度的阈值,且 $\beta_1 < \beta_2$,候选行为习惯簇的发生频率用符号 $\text{Fre}(HX_{p,q})$ 表示,候选行为习惯簇的密度用符号 $\text{Den}(HX_{p,q})$ 表示。

4. 如权利要求2所述的一种面向移动环境的学习行为习惯发现量化系统,其特征在于:所述的习惯发现单元通过对数据窗口中数据集进行聚类分析,将具有较高相似度的刺激情境和行为反应数据集合作为候选行为习惯簇 $HX_{x,y}$,当候选行为习惯簇的发生频率和密度同时达到某阈值时,认为此候选行为习惯簇是学习行为习惯簇,某候选行为习惯簇的发生频率和密度计算方法分别如下:

$$\text{Fre}(HX_{p,q}) = \frac{\text{count}(HX_{p,q})}{N_{total}} \quad (N_{total} \neq 0)$$

$$\text{Den}(HX_{p,q}) = \text{count}(HX_{p,q}) / t \quad (t \neq 0)$$

其中, $\text{count}(HX_{p,q})$ 为某候选行为习惯簇出现的次数,即该类簇中点的个数, N_{total} 为数据窗口中习惯特征数据的记录数, t 为第1次出现该候选行为习惯簇到最后1次出现该候选行为习惯簇之间的间隔天数,当 $\text{Fre}(HX_{p,q}) \leq \alpha_1 \cap \text{Den}(HX_{p,q}) \leq \beta_1$ 为真时,认为 $HX_{p,q}$ 不是学习行为习惯簇;当 $\text{Fre}(HX_{p,q}) \geq \alpha_2 \cap \text{Den}(HX_{p,q}) \geq \beta_2$ 为真时,认为 $HX_{p,q}$ 是学习行为习惯簇;当 $\alpha_1 < \text{Fre}(HX_{p,q}) < \alpha_2 \cap \beta_1 < \text{Den}(HX_{p,q}) < \beta_2$ 为真时,认为此时无法准确判断 $HX_{p,q}$ 是否为学习行为习惯簇,将基于时间的数据窗口向前扩大一个计算周期,再以 $HX_{p,q}$ 的聚类中心为新的聚类中心进行聚类、发生频率和密度的计算,当计算结果为 $\text{Fre}(HX_{p,q}) \leq \alpha_1 \cap \text{Den}(HX_{p,q}) \leq \beta_1$ 时,认为 $HX_{p,q}$ 不是学习行为习惯簇;当 $\text{Fre}(HX_{p,q}) \geq \alpha_2 \cap \text{Den}(HX_{p,q}) \geq \beta_2$ 为真时,认为 $HX_{p,q}$ 为学习行为习惯簇;当 $\alpha_1 < \text{Fre}(HX_{p,q}) < \alpha_2 \cap \beta_1 < \text{Den}(HX_{p,q}) < \beta_2$ 为真时,则将基于时间的数据窗口向后扩大一个计算周期,等下一个计算周期到达时,再以 $HX_{p,q}$ 的聚类中心为新的聚类中心进行聚类、发生频率和密度的计算,当计算结果为 $\text{Fre}(HX_{p,q}) \geq \alpha_2 \cap \text{Den}(HX_{p,q}) \geq \beta_2$ 时,认为 $HX_{p,q}$ 是学习行为习惯簇;当计算结果为 $\text{Fre}(HX_{p,q}) \leq \alpha_1 \cap \text{Den}(HX_{p,q}) \leq \beta_1$,或者 $\alpha_1 < \text{Fre}(HX_{p,q}) < \alpha_2 \cap \beta_1 < \text{Den}(HX_{p,q}) < \beta_2$,则认为 $HX_{p,q}$ 不是学习行为习惯簇,重新按默认情况下 $TW=3T$ 进行习惯发现计算。

5. 如权利要求2所述的一种面向移动环境的学习行为习惯发现量化系统,其特征就在于:所述的习惯量化单元量化学习者当前各学习行为习惯的强度时,根据赫尔学习理论对强化次数和习惯强度的函数关系的描述,量化学习行为习惯强度的计算公式如下:

$$HS_{x,y} = 1 - 10^{-0.0305 * count(H_{x,y})}$$

其中, $count(H_{x,y})$ 为数据窗口内强化学习行为习惯簇 $H_{x,y}$ 的次数,即数据窗口内该学习行为习惯簇发生的次数, $HS_{x,y}$ 为学习行为习惯强度。

6. 如权利要求1所述的一种面向移动环境的学习行为习惯发现量化系统,其特征就在于:所述移动学习终端与辅助穿戴设备通过蓝牙或无线局域网(WiFi)连接,所述移动学习终端与服务器通过无线局域网(WiFi)或公众移动通信网3G/4G/GPRS连接。

7. 如权利要求1所述的一种面向移动环境的学习行为习惯发现量化系统,其特征就在于:所述辅助穿戴设备包括但不限于智能手环、智能眼镜,所述移动学习终端包括但不限于手机、平板电脑。

8. 一种面向移动环境的学习行为习惯发现量化方法,其特征就在于,包括以下步骤:

(a) 利用辅助穿戴设备和移动学习终端收集学习者开展学习活动时的学习行为习惯关联数据;

(b) 经由移动学习终端的数据归集模块归集包含平台交互活动记录、学习者位置和生理状态的学习行为习惯关联数据,并将所述数据临时存储于数据存储模块中;

(c) 通过历史数据存储模块持久存储从移动学习终端历史数据存储模块上传至服务器的学习行为习惯关联数据;

(d) 利用数据预处理模块将历史数据存储模块中的学习行为习惯关联数据处理为便于机器识别和计算的习惯特征数据集;

(e) 利用聚类算法从习惯特征数据集中得到候选行为习惯簇,当无法准确判断某候选行为习惯簇是否为学习行为习惯簇时,利用数据控制单元控制数据窗口的大小,将调整数据窗口的大小,以进一步判断其是否为学习行为习惯簇;

(f) 经由习惯发现单元计算数据窗口中的习惯特征数据集,发现并获取学习者当前的学习行为习惯;

(g) 经由习惯量化单元,基于赫尔学习理论中强化次数与习惯强度之间的函数关系量化学习行为习惯;

(h) 通过习惯表征模块,向量化学习行为习惯簇,合并相似学习行为习惯的持续时长,从学习行为习惯描述,持续时长和强度三个维度表征学习行为习惯,并将学习行为习惯表征数据存储至历史数据存储模块;

(i) 利用移动学习终端的显示模块呈现习惯表征模块反馈的学习者学习行为习惯信息。

9. 如权利要求8所述的一种面向移动环境的学习行为习惯发现量化方法,其特征就在于,所述步骤(d)数据预处理模块的处理技术包括:

1) 数据变换,将数据转换成适合挖掘的形式;

2) 数据清洗,通过填写或忽略空缺值和错误数据、平滑噪声数据和孤立点、解决不一致性来清理数据;

3) 经过数据变换和数据清洗处理后,将平台交互数据表、位置信息表和生理状态信息

表中同一帐号的数据进行连接操作,得到一个习惯特征数据集X;
将习惯特征数据集X传入习惯计算模块。

一种面向移动环境的学习行为习惯发现量化系统及方法

技术领域

[0001] 本发明涉及移动学习技术领域,尤其涉及一种面向移动环境的学习行为习惯发现量化系统及方法。

背景技术

[0002] 随着智能设备的普及和移动互联网的推广,移动环境成为人们学习、生活和工作的主要场所,移动学习逐渐成为非正式学习的主要实施形式。移动环境为学习者提供了一个丰富便捷的新型学习形式,有利于多元化、多通道的信息互动,同时也在潜移默化中塑造学习者的学习行为习惯。良好学习行为习惯的养成,有助于激发学习者学习的积极性和主动性,培养自主学习能力,提高学习效率。移动计算环境支持下学习者佩戴或操作各类设备以开展自主学习、课堂学习等活动时,其平台交互活动记录、学习者位置和生理状态等信息将留存于移动设备及相关的服务端设备中,分析并计算这些数据以得到学习者的学习行为习惯,并将结果反馈给学习者,对提高自我管理能力,形成良好的移动学习行为习惯具有积极促进作用。

[0003] 目前关于网络学习行为习惯的研究并不多,面向移动学习环境的学习行为习惯研究更是寥寥数几,因此,为了顺应移动学习的趋势,研究如何在移动学习环境中即时掌握学习者的学习行为习惯,量化学习者的行为习惯强度显得尤为紧迫。

[0004] 学习习惯通常被认为是学习者在学习情境中通过反复实践养成的相对稳定的自动化行为模式。从习惯的形成来看,主要是由于一定的刺激情境与个体的某些动作在大脑皮层上形成稳固的暂时神经联系—条件反射链索系统。这样,当个体在同样的刺激情境下,条件反射的链索系统就会自动地出现,人就会自然而然地进行同样的有关动作。个体对刺激情境的反应包括两种,一是肉眼可观察到的动作,称为外显行为,二是肉眼不可观察到的动作,如心跳反应,称为内隐行为。因此,本发明认为移动学习环境中的一种学习行为要成为学习行为习惯,需要满足以下条件:1)固定刺激情境的出现;2)学习行为在该刺激情境中自动反复持续出现。因此,描述学习行为习惯的要素也应包含:1)刺激情境,包括学习者所处的物理环境和时间;2)行为反应,包括学习者的网络交互动作、对象、对象类型和心率波动等。

[0005] 传统的学习行为习惯获取忽略了移动学习环境中多通道数据蕴含的价值,采用的方法多为透过纸本或电子问卷的方式进行诊断,但是这种方法存在主观性过强、时效性有限的不足,此外,通过此方法也只能测量出问卷设计者提前设定的习惯,无法量化学习者学习行为习惯强度等问题。

发明内容

[0006] 有鉴于现有技术的上述缺陷,本发明所要解决的技术问题是提供一种面向移动环境的学习行为习惯发现量化系统及方法,针对传统方法主观性太强,时效性有限,只能测量出问卷设计者提前设定习惯等方面的缺点,本发明拟采用基于数据驱动的方法,实时收集

移动学习环境中学习者与平台交互的活动记录、所处位置和生理状态等外显和内隐的学习行为习惯关联数据,并将收集到的数据通过数据变换和数据清洗等方法处理为便于机器识别和计算的习惯特征数据集,进而基于聚类算法计算并发现学习行为习惯。针对传统方法无法量化学习者学习行为习惯强度的问题,本发明基于赫尔学习理论中强化次数与习惯强度之间的函数关系,量化学习行为习惯。最后,本发明从学习行为习惯的描述、持续时长和强度三个维度表征学习行为习惯。

[0007] 为实现上述目的,本发明提供了一种面向移动环境的学习行为习惯发现量化系统,系统由辅助穿戴设备,移动学习终端和服务端组成:

[0008] 辅助穿戴设备,由多个可监测学习者生理状态等信息的可穿戴设备构成,负责收集学习者开展学习活动时的生理状态信息;

[0009] 移动学习终端,移动学习终端宿居着一个习惯发现及量化系统的客户端,负责收集网络学习平台上的交互活动记录和学习者位置信息,归集与临时存储平台交互活动记录、学习者位置和生理状态的学习行为习惯关联数据,显示学习行为习惯的发现及量化结果,它包括如下三个模块:

[0010] 数据归集模块,它连接数据存储模块,用于归集从移动学习环境中收集到的平台交互活动记录、学习者位置和生理状态的学习行为习惯关联数据;

[0011] 数据存储模块,它连接数据归集模块和服务端中的历史数据存储模块,用于临时存储数据归集模块归集的学习行为习惯关联数据;

[0012] 显示模块,它连接服务端中的习惯表征模块,用于显示学习行为习惯的发现及量化结果;

[0013] 服务端,服务端上运行着一个习惯发现及量化系统的服务端,负责持久存储学习行为习惯关联数据和学习行为习惯表征数据,预处理与计算习惯关联数据以发现和量化学习者当前的学习行为习惯,它包括如下四个模块:

[0014] 历史数据存储模块,它连接数据存储模块、预处理模块和习惯表征模块,用于接收和持久存储移动学习终端数据存储模块传来的学习行为习惯关联数据,并传送至预处理模块,此外,还存储习惯表征模块的数据;

[0015] 预处理模块,它连接历史数据存储模块和习惯计算模块,用于将从历史数据存储模块中接收的学习行为习惯关联数据处理为便于机器识别和计算的 habit 特征数据集;

[0016] 习惯计算模块,它连接预处理模块和习惯表征模块,用于控制数据窗口的大小、计算并发现学习者的学习行为习惯簇、量化学习者当前各学习行为习惯的强度;

[0017] 习惯表征模块,它连接习惯计算模块、历史数据存储模块和显示模块,用于从学习行为习惯的描述、持续时长和强度三个维度表征学习行为习惯。

[0018] 进一步地,所述习惯计算模块还包括:

[0019] 数据窗口控制单元,它连接习惯发现单元,用于控制数据窗口的大小,当无法准确判断某候选行为习惯簇是否为学习行为习惯簇时,将调整数据窗口的大小,以进一步判断其是否为学习行为习惯簇;

[0020] 习惯发现单元,它连接数据窗口控制单元和习惯量化单元,用于计算数据窗口中的习惯特征数据,发现并获取学习者当前的学习行为习惯簇集合;

[0021] 习惯量化单元,它连接习惯发现单元,用于量化学习者当前各学习行为习惯的强

度。

[0022] 更进一步地,所述的数据窗口控制单元在习惯特征数据集X上基于时间段TW的数据窗口定义一个X的子集XW作为数据窗口的内容,TW为数据窗口的大小,TW满足以下公式:

$$[0023] \quad TW = \begin{cases} 3T, \\ 4T, \end{cases} \quad \alpha_1 > Fre(HX_{p,q}) > \alpha_2 \cap \beta_1 > Den(HX_{p,q}) > \beta_2$$

[0024] 默认情况下基于时间的数据窗口大小 $TW=3T$,但当 $\alpha_1 < Fre(HX_{p,q}) < \alpha_2 \cap \beta_1 < Den(HX_{p,q}) < \beta_2$ 时,认为此时无法准确判断候选行为习惯簇 $HX_{p,q}$ 是否为学习行为习惯簇,于是,将时间段TW向前扩大一个计算周期,使 $TW=4T$,对 $HX_{p,q}$ 再次进行判断:

[0025] 若计算结果为 $Fre(HX_{p,q}) \geq \alpha_2 \cap Den(HX_{p,q}) \geq \beta_2$,认为候选行为习惯簇 $HX_{p,q}$ 是学习行为习惯簇;

[0026] 若计算结果为 $Fre(HX_{p,q}) \leq \alpha_1 \cap Den(HX_{p,q}) \leq \beta_1$,认为候选行为习惯簇 $HX_{p,q}$ 不是学习行为习惯簇;

[0027] 若计算结果为 $\alpha_1 < Fre(HX_{p,q}) < \alpha_2 \cap \beta_1 < Den(HX_{p,q}) < \beta_2$,则将基于时间的数据窗口向后扩大一个计算周期,等下一个计算周期到达时,再对此候选行为习惯簇 $HX_{p,q}$ 进行习惯发现计算,

[0028] 若计算结果为 $Fre(HX_{p,q}) \geq \alpha_2 \cap Den(HX_{p,q}) \geq \beta_2$,认为候选行为习惯簇 $HX_{p,q}$ 是学习行为习惯簇;

[0029] 若计算结果为 $Fre(HX_{p,q}) \leq \alpha_1 \cap Den(HX_{p,q}) \leq \beta_1$,或者 $\alpha_1 < Fre(HX_{p,q}) < \alpha_2 \cap \beta_1 < Den(HX_{p,q}) < \beta_2$,则认为候选行为习惯簇 $HX_{p,q}$ 不是学习行为习惯簇,重新按默认情况下 $TW=3T$ 进行习惯发现计算。

[0030] 更进一步地,所述的习惯发现单元通过对数据窗口中数据集进行聚类分析,将具有较高相似度的刺激情境和行为反应数据集合作为候选行为习惯簇 $HX_{x,y}$,当候选行为习惯簇的发生频率和密度同时达到某阈值时,认为此候选行为习惯簇是学习行为习惯簇,某候选行为习惯簇的发生频率和密度计算方法分别如下:

$$[0031] \quad Fre(HX_{p,q}) = \frac{count(HX_{p,q})}{N_{total}} \quad (N_{total} \neq 0)$$

$$[0032] \quad Den(HX_{p,q}) = count(HX_{p,q}) / t \quad (t \neq 0)$$

[0033] 其中, $count(HX_{p,q})$ 为某候选行为习惯簇出现的次数,即该类簇中点的个数, N_{total} 为数据窗口中习惯特征数据的记录数, t 为第1次出现该候选行为习惯簇到最后1次出现该候选行为习惯簇之间的间隔天数,当 $Fre(HX_{p,q}) \leq \alpha_1 \cap Den(HX_{p,q}) \leq \beta_1$ 为真时,认为 $HX_{p,q}$ 不是学习行为习惯簇;当 $Fre(HX_{p,q}) \geq \alpha_2 \cap Den(HX_{p,q}) \geq \beta_2$ 为真时,认为 $HX_{p,q}$ 是学习行为习惯簇;当 $\alpha_1 < Fre(HX_{p,q}) < \alpha_2 \cap \beta_1 < Den(HX_{p,q}) < \beta_2$ 为真时,认为此时无法准确判断 $HX_{p,q}$ 是否为学习行为习惯簇,将基于时间的数据窗口向前扩大一个计算周期,再以 $HX_{p,q}$ 的聚类中心为新的聚类中心进行聚类、发生频率和密度的计算,当计算结果为 $Fre(HX_{p,q}) \leq \alpha_1 \cap Den(HX_{p,q}) \leq \beta_1$ 时,认为 $HX_{p,q}$ 不是学习行为习惯簇;当 $Fre(HX_{p,q}) \geq \alpha_2 \cap Den(HX_{p,q}) \geq \beta_2$ 为真时,认为 $HX_{p,q}$ 为学习行为习惯簇;当 $\alpha_1 < Fre(HX_{p,q}) < \alpha_2 \cap \beta_1 < Den(HX_{p,q}) < \beta_2$ 为真时,则将基于时间的数据窗口向后扩大一个计算周期,等下一个计算周期到达时,再以 $HX_{p,q}$ 的聚类中心为新的聚类中心进行聚类、发生频率和密度的计算,当计算结果为 $Fre(HX_{p,q}) \geq \alpha_2 \cap Den(HX_{p,q}) \geq \beta_2$ 时,认为 $HX_{p,q}$ 是学习行为习惯簇;当计算结果为 $Fre(HX_{p,q}) \leq \alpha_1 \cap Den(HX_{p,q}) \leq \beta_1$,或者 $\alpha_1 < Fre(HX_{p,q})$

$\langle \alpha_2 \cap \beta_1 \langle \text{Den}(\text{HX}_{p,q}) \rangle \langle \beta_2 \rangle$, 则认为 $\text{HX}_{p,q}$ 不是学习行为习惯簇, 重新按默认情况下 $\text{TW} = 3\text{T}$ 进行习惯发现计算。

[0034] 更进一步地, 所述的习惯量化单元量化学习者当前各学习行为习惯的强度时, 根据赫尔学习理论对强化次数和习惯强度的函数关系的描述, 量化学习行为习惯强度的计算公式如下:

$$[0035] \quad \text{HS}_{x,y} = 1 - 10^{-0.0305 * \text{count}(\text{H}_{x,y})}$$

[0036] 其中, $\text{count}(\text{H}_{x,y})$ 为数据窗口内强化学习行为习惯簇 $\text{H}_{x,y}$ 的次数, 即数据窗口内该学习行为习惯簇发生的次数, $\text{HS}_{x,y}$ 为学习行为习惯强度。

[0037] 进一步地, 所述移动学习终端与辅助穿戴设备通过蓝牙或无线局域网 (WiFi) 连接, 所述移动学习终端与服务器通过无线局域网 (WiFi) 或公众移动通信网 3G/4G/GPRS 连接。

[0038] 进一步地, 所述辅助穿戴设备包括但不限于智能手环、智能眼镜, 所述移动学习终端包括但不限于手机、平板电脑。

[0039] 一种面向移动环境的学习行为习惯发现量化方法, 包括以下步骤:

[0040] (a) 利用辅助穿戴设备和移动学习终端收集学习者开展学习活动时的学习行为习惯关联数据;

[0041] (b) 经由移动学习终端的数据归集模块归集包含平台交互活动记录、学习者位置和生理状态的学习行为习惯关联数据, 并将所述数据临时存储于数据存储模块中;

[0042] (c) 通过历史数据存储模块持久存储从移动学习终端历史数据存储模块上传至服务器的学习行为习惯关联数据;

[0043] (d) 利用数据预处理模块将历史数据存储模块中的学习行为习惯关联数据处理为便于机器识别和计算的习惯特征数据集;

[0044] (e) 利用聚类算法从习惯特征数据集中得到候选行为习惯簇, 当无法准确判断某候选行为习惯簇是否为学习行为习惯簇时, 利用数据控制单元控制数据窗口的大小, 将调整数据窗口的大小, 以进一步判断其是否为学习行为习惯簇;

[0045] (f) 经由习惯发现单元计算数据窗口中的习惯特征数据集, 发现并获取学习者当前的学习行为习惯;

[0046] (g) 经由习惯量化单元, 基于赫尔学习理论中强化次数与习惯强度之间的函数关系量化学习行为习惯;

[0047] (h) 通过习惯表征模块, 向量化学习行为习惯簇, 合并相似学习行为习惯的持续时长, 从学习行为习惯描述, 持续时长和强度三个维度表征学习行为习惯, 并将学习行为习惯表征数据存储至历史数据存储模块;

[0048] (i) 利用移动学习终端的显示模块呈现习惯表征模块反馈的学习者学习行为习惯信息。

[0049] 进一步地, 所述步骤 (d) 数据预处理模块的处理技术包括:

[0050] 1) 数据变换, 将数据转换成适合挖掘的形式;

[0051] 2) 数据清洗, 通过填写或忽略空缺值和错误数据、平滑噪声数据和孤立点、解决不一致性来清理数据;

[0052] 3) 经过数据变换和数据清洗处理后, 将平台交互数据表、位置信息表和生理状态

信息表中同一帐号的数据进行连接操作,得到一个习惯特征数据集X;

[0053] 4) 将习惯特征数据集X传入习惯计算模块。

[0054] 本发明的有益效果是:

[0055] 本发明通过收集、预处理和计算移动学习环境中包含平台交互活动记录、学习者位置和生理状态等外显和内隐的学习行为习惯关联数据,可以在线自动发现和量化学习者的学习行为习惯。其次,在学习行为习惯特征信息的收集中考虑学习者所处位置和生理状态信息,多维度的信息来源能更客观准确地反映学习者的学习行为习惯。再者,提出的一种面向移动学习环境的学习行为习惯发现及量化系统和方法可根据对候选行为习惯簇的计算,控制用于学习行为习惯发现与量化计算时输入数据的大小,使之能更准确地发现学习行为习惯。

[0056] 以下将结合附图对本发明的构思、具体结构及产生的技术效果作进一步说明,以充分地了解本发明的目的、特征和效果。

附图说明

[0057] 图1是本发明的系统结构图。

[0058] 图2是本发明的发现学习行为习惯簇的具体流程图。

[0059] 图3是本发明的发现学习行为习惯簇的具体流程图。

具体实施方式

[0060] 本发明的一种面向移动环境的学习行为习惯发现量化系统结构图如图1所示,系统由辅助穿戴设备,移动学习终端和服务端组成。

[0061] 辅助穿戴设备:由多个可监测学习者生理状态等信息的可穿戴设备构成,负责收集学习者开展学习活动时的生理状态等信息。

[0062] 移动学习终端:移动学习终端宿居着一个习惯发现及量化系统的客户端,负责收集网络学习平台上的交互活动记录和学习者位置信息,归集与临时存储平台交互活动记录、学习者位置和生理状态等学习行为习惯关联数据,显示学习行为习惯的发现及量化结果。它包括如下三个模块:

[0063] 数据归集模块,它连接数据存储模块,用于归集从移动学习环境中收集到的平台交互活动记录、学习者位置和生理状态等学习行为习惯关联数据;

[0064] 数据存储模块,它连接数据归集模块和服务端中的历史数据存储模块,用于临时存储数据归集模块归集的学习行为习惯关联数据;

[0065] 显示模块,它连接服务端中的习惯表征模块,用于显示学习行为习惯的发现及量化结果。

[0066] 服务器:服务器上运行着一个习惯发现及量化系统的服务端,负责持久存储学习行为习惯关联数据和学习行为习惯表征数据,预处理与计算习惯关联数据以发现和量化学习者当前的学习行为习惯。它包括如下四个模块:

[0067] 历史数据存储模块,它连接数据存储模块、预处理模块和习惯表征模块,用于接收和持久存储移动学习终端数据存储模块传来的学习行为习惯关联数据,并传送至预处理模块,此外,还存储习惯表征模块的数据;

[0068] 预处理模块,它连接历史数据存储模块和习惯计算模块,用于将从历史数据存储模块中接收的学习行为习惯关联数据处理为便于机器识别和计算的习惯特征数据集;

[0069] 习惯计算模块,它连接预处理模块和习惯表征模块,用于控制数据窗口的大小、计算并发现学习者的学习行为习惯簇、量化学习者当前各学习行为习惯的强度;

[0070] 习惯表征模块,它连接习惯计算模块、历史数据存储模块和显示模块,用于从学习行为习惯的描述,持续时长和强度三个维度表征学习行为习惯。

[0071] 其中,习惯计算模块还包括:

[0072] 数据窗口控制单元,它连接习惯发现单元,用于控制数据窗口的大小,当无法准确判断某候选行为习惯簇是否为学习行为习惯簇时,将调整数据窗口的大小,以进一步判断其是否为学习行为习惯簇;

[0073] 习惯发现单元,它连接数据窗口控制单元和习惯量化单元,用于计算数据窗口中的习惯特征数据,发现并获取学习者当前的学习行为习惯簇集合;

[0074] 习惯量化单元,它连接习惯发现单元,用于量化学习者当前各学习行为习惯的强度。

[0075] 在实施中,移动学习终端与辅助穿戴设备可通过蓝牙或无线局域网(WiFi)连接,移动学习终端与服务器通过无线局域网(WiFi)或公众移动通信网(如4G,3G或GPRS)连接。辅助穿戴设备可以是智能手环、智能眼镜等包含多种传感器的可穿戴设备,移动学习终端可以为手机、平板电脑等。

[0076] 本发明所提供的一种面向移动学习环境的学习行为习惯量化方法包含以下步骤:

[0077] (a) 利用辅助穿戴设备和移动学习终端收集学习者开展学习活动时的学习行为习惯关联数据;

[0078] (b) 经由移动学习终端的数据归集模块归集包含平台交互活动记录、学习者位置和生理状态等的学习行为习惯关联数据,并将所述数据临时存储于数据存储模块中;

[0079] (c) 通过历史数据存储模块持久存储从移动学习终端历史数据存储模块上传至服务器的学习行为习惯关联数据;

[0080] (d) 利用数据预处理模块将历史数据存储模块中的学习行为习惯关联数据处理为便于机器识别和计算的 habit 特征数据集;

[0081] (e) 利用聚类算法从习惯特征数据集中得到候选行为习惯簇,当无法准确判断某候选行为习惯簇是否为学习行为习惯簇时,利用数据控制单元控制数据窗口的大小,将调整数据窗口的大小,以进一步判断其是否为学习行为习惯簇;

[0082] (f) 经由习惯发现单元计算数据窗口中的习惯特征数据集,发现并获取学习者当前的学习行为习惯;

[0083] (g) 经由习惯量化单元,基于赫尔学习理论中强化次数与习惯强度之间的函数关系量化学习行为习惯;

[0084] (h) 通过习惯表征模块,向量化学习行为习惯簇,合并相似学习行为习惯的持续时长,从学习行为习惯描述,持续时长和强度三个维度表征学习行为习惯,并将学习行为习惯表征数据存储至历史数据存储模块;

[0085] (i) 利用移动学习终端的显示模块呈现习惯表征模块反馈的学习者学习行为习惯信息;

[0086] 下面,结合附图与具体实施例对本发明作进一步说明,为了更清楚说明面向移动学习环境的学习行为习惯的发现及量化过程,对相关概念做了如下定义:

[0087] 定义1:习惯特征数据集

[0088] 习惯特征数据集是指通过数据预处理后得到平台交互活动记录、学习者位置和生理状态的一系列数据集。记作 $X, X = \{x_1, x_2, \dots, x_i, \dots, x_n\}, n \in \mathbb{N}^*$, 其中, $x_i = (x_{i1}, x_{i2}, \dots, x_{ik}) \in \mathbb{R}^k, k$ 表示数据对象的空间维数。

[0089] 定义2:计算周期

[0090] 计算周期,为了持续关注学习者的学习行为习惯动态,每隔一段时间将进行学习行为习惯的发现和量化计算,这段时间长度即为计算周期,记作 T 。由于生活中一般以7天为一个学习周期,因此本实施例中设 $T=7$ 天。

[0091] 定义3:候选行为习惯簇

[0092] 候选行为习惯簇是指对数据窗口中习惯特征数据集进行聚类后得到的具有较高相似度的刺激情境和行为反应的数据集,亦即聚类结果中的某个类簇,将第 p 次计算得到的第 q 个候选行为习惯簇记作 $HX_{p,q}, HX_{p,q} = \{x_1', x_2', \dots, x_i', \dots, x_n'\}, n, p, q \in \mathbb{N}^*, x_i' \in X$ 。

[0093] 定义4:候选行为习惯簇的发生频率

[0094] 候选行为习惯簇的发生频率,指数据窗口内候选行为习惯簇发生的次数与习惯特征数据集的记录数之比,是描述学习者候选行为习惯簇发生频繁程度的量,用符号 $Fre(HX_{p,q})$ 表示,定义 α_1, α_2 为候选行为习惯簇发生频率的阈值,且 $\alpha_1 < \alpha_2$ 。

[0095] 定义5:候选行为习惯簇的密度

[0096] 候选行为习惯簇的密度,指单位时间内候选行为习惯簇出现的次数,是描述学习者候选行为习惯簇出现密集程度的量,用符号 $Den(HX_{p,q})$ 表示,定义 β_1, β_2 为候选行为习惯簇密度的阈值,且 $\beta_1 < \beta_2$ 。

[0097] 定义6:学习行为习惯簇

[0098] 学习行为习惯簇是学习者在某种固定刺激情境中反复持续出现的一种相对稳定的学习行为数据集,即达到一定发生频率和密度的候选行为习惯簇。将第 x 次计算得到的第 y 个学习行为习惯簇记作 $H_{x,y}, H_{x,y} = \{x_1'', x_2'', \dots, x_i'', \dots, x_n''\}, x, y, n \in \mathbb{N}^*, x_i'' \in X^*$ 。当 $Fre(HX_{p,q}) \geq \alpha_2 \cap Den(HX_{p,q}) \geq \beta_2$ 为真时,认为候选行为习惯簇 $HX_{p,q}$ 为学习行为习惯簇,当 $\alpha_1 < Fre(HX_{p,q}) < \alpha_2 \cap \beta_1 < Den(HX_{p,q}) < \beta_2$ 为真时,认为此时无法准确判断该候选行为习惯簇 $HX_{p,q}$ 是否为学习行为习惯簇,当 $Fre(HX_{p,q}) \leq \alpha_1 \cap Den(HX_{p,q}) \leq \beta_1$ 为真时,认为候选行为习惯簇 $HX_{p,q}$ 不是学习行为习惯簇。

[0099] 定义7:学习行为习惯表征

[0100] 学习行为习惯表征是指用学习行为习惯的描述、持续时长、强度三个指标表征学习行为习惯,其中,学习行为习惯的描述可以从起始时间、终止时间、场所、网络交互动作、对象、对象类型和心率波动七方面展开。

[0101] 定义8:数据窗口

[0102] 数据窗口是一种控制习惯发现时输入数据范围的控件,当在时刻 t_i 进行习惯发现计算时,习惯特征数据集 X 上基于时间段 TW 的数据窗口定义了一个 X 的子集 XW ,子集 XW 中所有数据记录的时刻 t_i' 满足 $t_i - t_i' \leq TW$,即给定习惯特征数据集 X ,基于时间的数据窗口大小为 TW ,时间段 $[t_i - TW, t_i]$ 内 X 的子集 XW 作为数据窗口的内容。 TW 满足公式(2):

$$TW = \begin{cases} 3T, & \text{default} \\ 4T, & \alpha_1 < \text{Fre}(\mathbf{HX}_{p,q}) < \alpha_2 \cap \beta_1 < \text{Der}(\mathbf{HX}_{p,q}) < \beta_2 \end{cases} \quad (2)$$

[0104] 根据以上系统、方法和相关定义,接下来以一个具体实施例进行阐述。

[0105] 学习者在移动学习环境中开展自主学习、课堂学习等活动时,通过手机或平板电脑等移动学习终端设备进行人机交互,其在学习平台中的交互数据及位置信息可被该设备收集,同时,学习者在使用可穿戴设备时,其生理状态数据将被收集到可穿戴设备中。由于需要实时收集学习者的生理数据,因此,可穿戴设备和移动学习终端以10秒为采样周期收集学习者开展学习活动时的平台交互活动记录、位置和生理状态三类学习行为习惯关联数据,数据归集模块则基于蓝牙或无线局域网技术归集可穿戴设备和移动学习终端中收集到的三类学习行为习惯关联数据,并将这些数据分别存储至移动学习终端数据存储模块的以下三个表中:

[0106] 1) 平台交互数据表(属性包括:编号、学习者编号、时间、网络交互动作、对象及对象类型)

[0107] 2) 位置信息表(属性包括:编号、学习者编号、时间、位置和场所)

[0108] 3) 生理状态信息表(属性包括:编号、学习者编号、时间、心率波动)。

[0109] 其中,心率波动HF用当前所测得的心率值 z_i 与正常人的平均心率值 μ 之差表示,计算方法如下:

$$HF = z_i - \mu \quad (z_i, \mu \in \mathbb{N}^*) \quad (2)$$

[0111] 其中,位置信息表的场所值由移动设备接收到的POI信息所包含的位置类型决定。POI是“Point of Interest”的缩写,可翻译为“兴趣点”,每个POI包含四方面信息:名称、位置类别、经度、纬度。当使用地图软件定位时,将返回所处位置的POI信息。当到达预设的学习行为习惯计算周期时,学习行为习惯关联数据将被上传至服务器,若到达规定上传时间时移动学习终端未联网或发生其他故障导致数据未被上传至服务器,则当移动学习终端再次联网或恢复正常时立刻上传数据至服务器。数据被上传至服务器后,将被存储于历史数据存储模块。接着对历史数据存储模块中的学习行为习惯关联数据进行预处理,数据预处理技术包括:

[0112] 数据变换:将数据转换成适合挖掘的形式。离散化时间属性,将时间用时间序号表示,对00:00:00-24:00:00中的每个10秒进行标号,即00:00:00-00:00:10,时间序号为1,00:00:10-00:00:20,时间序号为2,以此类推。此外,对场所、动作、对象及对象类型属性进行数字化处理。将场所信息按平台交互数据表所示的场所信息数字化对照表数字化,场所的数字编号由一级场所分类的编号和二级场所分类的编号组合而成,例如:居住场所一宿舍的编号为0202。将动作、对象及对象类型按位置信息表位置信息表所示的平台交互数据数字化对照表进行数字化,分别将动作属性中的浏览、发布、练习、使用学习辅助工具和交流数字化为01、02、03、04、05,分别将对象中的课件、笔记、帮助文档、日志、文章、消息、提问、投票、评论、习题、试卷、导航、搜索引擎、收藏夹、好友动态、论坛、小组活动数字化为01至17,分别将对象类型中的文字、图片、图表、声音、视频、动画数字化为01、02、03、04、05、06。

[0113] 数据清洗:通过填写或忽略空缺值和错误数据、平滑噪声数据和孤立点、解决不一致性来清理数据。例如,表中某条记录的心率值为276,该值严重偏离了心率的正常范围,记

录出错,则重新填写一个缺省值,缺省值为心率记录中上一条记录与下一条记录心率值的平均值;属性值空缺的填补办法为:空缺的场所值由上一条记录的场所值进行填充,空缺的心率波动值则由缺省值填充。

[0114] 经过数据变换和数据清洗处理后,将平台交互数据表、位置信息表和生理状态信息表中同一帐号的数据进行连接操作,得到一个习惯特征数据集X。连接方法为:首先对生理状态信息表与平台交互数据表进行全外连接操作,再将得到的结果与位置信息表进行全外连接操作,最终得到融合三个数据表的数据集X。在本实施例中学习行为习惯特征由起始时间序号、终止时间序号、场所、网络交互动作、对象、对象类型和心率波动七元组 $x_i = (x_{i1}, x_{i2}, x_{i3}, x_{i4}, x_{i5}, x_{i6}, x_{i7})$ 构成,每条数据记录可以描述某学习者在某时某地对具有某种类型的对象进行某种网络交互动作,并记录该状态下的心率波动。生理状态信息表展示了某学习者习惯特征数据集中的部分数据。其中,时间序号3241,3242,3243分别代表09:00:00-09:00:10,09:00:10-09:00:20,09:00:20-09:00:30三个时间段,场所0202代表居住场所-宿舍,动作01代表浏览,对象01代表课件,对象类型02代表图片,时间序号为3241的这条数据记录学习者09:00:00-09:00:10在宿舍利用移动学习平台浏览课件的学习行为,该课件类型为图片,此时学习者心率波动值为7。

[0115] 接着将习惯特征数据集X传入习惯计算模块。首先利用数据窗口控制单元来控制数据窗口的大小,本实施例中采用基于时间的数据窗口以确定位于窗口内的数据点。

[0116] 默认情况下基于时间的数据窗口大小 $TW = 3T$,然后进行学习行为习惯簇发现计算,当 $\alpha_1 < Fre(HX_{p,q}) < \alpha_2 \cap \beta_1 < Den(HX_{p,q}) < \beta_2$ 时, $TW = 4T$ (向前加T),若计算结果仍为 $\alpha_1 < Fre(HX_{p,q}) < \alpha_2 \cap \beta_1 < Den(HX_{p,q}) < \beta_2$, $TW = 4T$ (向后加T),再次进行习惯发现计算,若计算结果仍为 $\alpha_1 < Fre(HX_{p,q}) < \alpha_2 \cap \beta_1 < Den(HX_{p,q}) < \beta_2$,则认为该候选行为习惯簇不是学习行为习惯簇,令 $TW = 3T$ 继续计算。

[0117] 学习行为习惯发现计算的基本思想为:首先利用聚类算法对数据窗口内的习惯特征数据集XW进行聚类,得到多个候选行为习惯簇 $HX_{p,q}$,然后计算某候选行为习惯簇的发生频率和密度,当 $Fre(HX_{p,q}) \geq \alpha_2 \cap Den(HX_{p,q}) \geq \beta_2$ 为真时,认为候选行为习惯簇 $HX_{p,q}$ 为学习行为习惯簇;当 $\alpha_1 < Fre(HX_{p,q}) < \alpha_2 \cap \beta_1 < Den(HX_{p,q}) < \beta_2$ 为真时,认为此时无法准确判断候选行为习惯簇 $HX_{p,q}$ 是否为学习行为习惯簇,于是,将数据窗口的大小向前或向后扩大一个计算周期,然后再针对此候选行为习惯簇进行习惯发现计算;当 $Fre(HX_{p,q}) \leq \alpha_1 \cap Den(HX_{p,q}) \leq \beta_1$ 为真时,认为候选行为习惯簇 $HX_{p,q}$ 不是学习行为习惯簇。对所有候选行为习惯簇进行以上计算,最终得到该学习者当前的学习行为习惯簇的集合。

[0118] 本实施例采用一种密度峰值快速搜索聚类算法(DPC),该算法的基本原理:理想的类簇中心具备两个基本特征:1)其局部密度大于围绕它邻居的局部密度;2)不同类簇中心之间的距离相对较远。为了找到同时满足上述条件的类簇中心,DPC算法引入了点 x_i 的局部密度 ρ_i 和点 x_i 到局部密度比它大且距离它最近的点的距离 δ_i ,在本实施例中, x_i 和 x_j 为数据窗口中习惯特征数据集的任意两条记录。为了更清楚地描述DPC算法的实施过程,对任意两点间距离、局部密度等概念进行定义,具体如下:

[0119] 定义9任意两点间距离:用于计算任意两点间的距离,公式如下:

$$[0120] \quad d_{ij} = \sqrt{\lambda_1(|x_{i1} - x_{j1}|)^2 + \lambda_2(|x_{i2} - x_{j2}|)^2 + \dots + \lambda_k(|x_{ik} - x_{jk}|)^2} \quad (3)$$

[0121] 其中 $x_i = (x_{i1}, x_{i2}, \dots, x_{ik}) \in R^k$, k 为数据对象的空间维数, $\lambda_1, \lambda_2, \dots, \lambda_k \in (0, 1)$ 为权重系数, 并且满足 $\lambda_1 + \lambda_2 + \dots + \lambda_k = 1$, 根据习惯特征数据集中各属性对学习行为习惯的影响力大小确定。

[0122] 定义10局部密度: 用于计算与点 x_i 的距离小于截断距离的点的个数 (不包含点 x_i 本身), 公式如下:

$$[0123] \quad \rho_i = \sum_j \chi(d_{ij} - d_c) \quad (4)$$

[0124] 其中, $\chi(x) = \begin{cases} 1 & \text{if } x < 0, \\ 0 & \text{otherwise} \end{cases}$, d_{ij} 为任意两点 x_i 和 x_j 间的欧式距离, d_c 是截断距离。

[0125] 定义11高局部密度点距离: 用于计算点 x_i 到局部密度大于点 x_i 且离它最近的点 x_j 的距离, 公式如下:

$$[0126] \quad \delta_i = \min_{j: \rho_j > \rho_i} (d_{ij}) \quad (5)$$

[0127] 对于局部密度 ρ_i 最大的点 x_i , 其 $\delta_i = \max_j (d_{ij})$ 。

[0128] DPC算法的描述如下:

[0129] 输入: 数据窗口内的习惯特征数据集 XW ;

[0130] 输出: 类簇集合 CHX ;

[0131] Begin:

[0132] Step1: 根据式 (3) 计算任意两点 x_i 和 x_j 之间的距离 d_{ij} ;

[0133] Step2: 将所有点的相互距离 d_{ij} 进行升序排列, 占总数 2% 位置的距离值定义为 d_c 的值;

[0134] Step3: 根据式 (4) 计算得到每个点的局部密度 ρ_i ;

[0135] Step4: 根据式 (5) 计算得到每个点的 δ_i 值;

[0136] Step5: 将具有较大距离 δ_i 且同时具有较大局部密度 ρ_i 的点定义为聚类中心;

[0137] Step6: 将剩余的每个点归属到与其最邻近且密度比其大的节点所属的类簇中, 不断进行迭代直到没有剩余点;

[0138] Step7: 将高局部密度点距离 δ_i 较大但局部密度 ρ_i 较小的孤立点去除;

[0139] Step8: 输出类簇集合 CHX ;

[0140] End

[0141] 最终得到类簇集合 CHX , 即候选行为习惯簇的集合, 任意候选行为习惯簇 $HX_{p,q} \in CHX$ 。接着进行候选行为习惯簇发生频率的计算: 定义某候选行为习惯簇发生的次数, 即该候选行为习惯簇 $HX_{p,q}$ 中记录的条目数为 $\text{count}(HX_{p,q})$, 数据窗口中习惯特征数据集的记录数为 N_{total} , 某候选行为习惯簇的发生频率计算方法如公式 (6):

$$[0142] \quad \text{Fre}(HX_{p,q}) = \text{count}(HX_{p,q}) / N_{\text{total}} \quad (N_{\text{total}} \neq 0) \quad (6)$$

[0143] 候选行为习惯簇密度的计算: 定义第 1 次出现某候选行为习惯簇到最后 1 次出现该候选行为习惯簇之间的间隔天数为 t , 学习者某候选行为习惯簇的密度计算方法如公式 (7):

$$[0144] \quad \text{Den}(HX_{p,q}) = \text{count}(HX_{p,q}) / t \quad (t \neq 0) \quad (7)$$

[0145] 当 $\text{Fre}(HX_{p,q}) \leq \alpha_1 \cap \text{Den}(HX_{p,q}) \leq \beta_1$ 为真时, 认为 $HX_{p,q}$ 不是学习行为习惯簇; 当 Fre

$(HX_{p,q}) \geq \alpha_2 \cap \text{Den}(HX_{p,q}) \geq \beta_2$ 为真时,认为 $HX_{p,q}$ 是学习行为习惯簇;当 $\alpha_1 < \text{Fre}(HX_{p,q}) < \alpha_2 \cap \beta_1 < \text{Den}(HX_{p,q}) < \beta_2$ 为真时,认为此时无法准确判断 $HX_{p,q}$ 是否为学习行为习惯簇,于是,将基于时间的数据窗口向前扩大一个计算周期,再以 $HX_{p,q}$ 的聚类中心为新的聚类中心进行聚类、发生频率和密度的计算,当计算结果为 $\text{Fre}(HX_{p,q}) \leq \alpha_1 \cap \text{Den}(HX_{p,q}) \leq \beta_1$ 时,认为 $HX_{p,q}$ 不是学习行为习惯簇;当 $\text{Fre}(HX_{p,q}) \geq \alpha_2 \cap \text{Den}(HX_{p,q}) \geq \beta_2$ 为真时,认为 $HX_{p,q}$ 为学习行为习惯簇;当 $\alpha_1 < \text{Fre}(HX_{p,q}) < \alpha_2 \cap \beta_1 < \text{Den}(HX_{p,q}) < \beta_2$ 为真时,则将基于时间的数据窗口向后扩大一个计算周期,等下一个计算周期到达时,再以 $HX_{p,q}$ 的聚类中心为新的聚类中心进行聚类、发生频率和密度的计算,当计算结果仍为 $\text{Fre}(HX_{p,q}) \geq \alpha_2 \cap \text{Den}(HX_{p,q}) \geq \beta_2$ 时,认为 $HX_{p,q}$ 是学习行为习惯簇;当计算结果为 $\text{Fre}(HX_{p,q}) \leq \alpha_1 \cap \text{Den}(HX_{p,q}) \leq \beta_1$,或者 $\alpha_1 < \text{Fre}(HX_{p,q}) < \alpha_2 \cap \beta_1 < \text{Den}(HX_{p,q}) < \beta_2$,则认为 $HX_{p,q}$ 不是学习行为习惯簇,重新按默认情况下 $TW=3T$ 进行习惯发现计算。

[0146] 对所有候选行为习惯簇进行以上计算,最终得到学习者当前的学习行为习惯簇集合,此时得到的学习行为习惯簇 $H_{x,y}$ 是由多条相似度极高的习惯特征数据记录组成的数据集。为了方便下文描述学习行为习惯,本实施例中学习行为习惯簇 $H_{x,y}$ 转化为七元组 $H_{x,y} = (a_1, a_2, a_3, a_4, a_5, a_6, a_7)$, a_1 表示起始时间序号, a_2 表示终止时间序号, a_3 表示场所, a_4 表示动作, a_5 表示对象, a_6 表示对象类型, a_7 表示心率波动。发现学习行为习惯簇的计算流程如图2和图3所示。

[0147] 最后利用习惯量化单元量化学习者当前各学习行为习惯的强度,根据赫尔学习理论对强化次数和习惯强度的函数关系的描述,量化学习行为习惯强度的计算公式如下:

$$[0148] \quad HS_{x,y} = 1 - 10^{-0.0305 * \text{count}(H_{x,y})} \quad (8)$$

[0149] 其中, $\text{count}(H_{x,y})$ 为数据窗口内强化学习行为习惯簇 $H_{x,y}$ 的次数,即数据窗口内该学习行为习惯簇发生的次数, $HS_{x,y}$ 为学习行为习惯强度。

[0150] 习惯计算模块完成所有计算后,将结果传至习惯表征模块。该模块从学习行为习惯的描述,持续时长和强度三个维度表征学习行为习惯,第 x 次计算得到的第 y 个学习行为习惯表征可以用一个三元组 $\text{Cha}H_{x,y} = (\text{vec}H_{x,y}, \text{Dur}_{x,y}, HS_{x,y})$ 表示。其中,学习行为习惯强度已在习惯量化单元计算得出,对学习行为习惯的描述和持续时长则需进一步的阐述。

[0151] 为方便对学习行为习惯进行描述,将从起始时间、终止时间、场所、网络交互动作、对象、对象类型和心率波动七个维度对学习行为习惯发现单元得到的学习行为习惯簇 $H_{x,y}$ 进行向量化处理,将第 x 次计算得到的第 y 个向量化后的学习行为习惯记作 $\text{vec}H_{x,y} = (v_{y1}, v_{y2}, v_{y3}, v_{y4}, v_{y5}, v_{y6}, v_{y7})$,其中, $v_{y3}, v_{y4}, v_{y5}, v_{y6}, v_{y7}$ 各属性的值由第 y 个学习行为习惯簇 $H_{x,y}$ 中与之对应的属性值众数表示,定义 $a_i | x_j$ ”表示从学习行为习惯簇 $H_{x,y}$ 中任意数据对象 x_j ”中取出属性 a_i 的值,定义 $M_0(a_i | x_j)$ ”表示学习行为习惯簇 $H_{x,y}$ 中属性 a_i 的值的众数,那么有 $\text{value}(v_{yi}) = M_0(a_{i-1} | x_j)$, $i \in \{3, 4, 5, 6, 7\}$,起始时间和终止时间维度则是将学习行为习惯簇 $H_{x,y}$ 中最小的时间序号属性值作为起始时间,最大值作为终止时间,那么有 $\text{value}(v_{y1}) = \min(a_1 | x_j)$, $\text{value}(v_{y2}) = \max(a_1 | x_j)$ ”。

[0152] 学习行为习惯的持续时长计算方法为:将最新计算得到的向量化学习行为习惯集合,即第 M 次计算得到的向量化学习行为习惯集合 $\text{Vec}H_M = (\text{vec}H_{M,1}, \text{vec}H_{M,2}, \dots, \text{vec}H_{M,n})$ 中的每个向量化学习行为习惯 $\text{vec}H_{M,y} = (v_{y1}, v_{y2}, \dots, v_{y7})$,与第 $M-1$ 次计算得到的向量化学习行为习惯集合 $\text{Vec}H_{M-1} = (\text{vec}H_{M-1,1}, \text{vec}H_{M-1,2}, \dots, \text{vec}H_{M-1,n'})$ 中的每个向量化学习行为习惯 $\text{vec}H_{M-1,y'} = (v_{y'1}, v_{y'2}, \dots, v_{y'7})$ 进行相似度计算,若相似度大于或等于阈值 η ,则认为这两

个向量化学习行为习惯是一样的,那么第M次计算得到的第y个学习行为习惯的持续时长 $Dur_{M,y}$ 为第M-1次计算得到的第y'个学习行为习惯的持续时长 $Dur_{M-1,y'}$ 与计算周期T之和,若相似度小于阈值 η ,那么第M次计算得到的第y个学习行为习惯的持续时长 $Dur_{M,y}$ 为计算该学习行为习惯时数据窗口的大小,计算公式如下:

$$[0153] \quad \begin{aligned} HSI(\mathit{vecH}_M, \mathit{vecH}_{M-1,y'}) &= \text{sim}(\mathit{vecH}_{M,y}, \mathit{vecH}_{M-1,y'}) \\ &= \frac{\sum_{i=1}^7 (v_{yi} \times v_{y'i})}{\left(\sqrt{\sum_{i=1}^7 (v_{yi}^2)} \times \sqrt{\sum_{i=1}^7 (v_{y'i}^2)}\right)} \end{aligned} \quad (9)$$

$$[0154] \quad Dur_{M,y} = \begin{cases} Dur_{M-1,y'} + T, & HSI(\mathit{vecH}_M, \mathit{vecH}_{M-1,y'}) \geq \eta \\ TW, & HSI(\mathit{vecH}_M, \mathit{vecH}_{M-1,y'}) < \eta \end{cases} \quad (10)$$

[0155] 习惯表征模块从学习行为习惯的描述,持续时长和强度三个维度表征学习行为习惯后,将结果作为学习行为习惯表征数据存储至历史数据存储模块,同时还将结果传至移动学习终端的显示模块。

[0156] 以上详细描述了本发明的较佳具体实施例。应当理解,本领域的普通技术人员无需创造性劳动就可以根据本发明的构思做出诸多修改和变化。因此,凡本技术领域技术人员依本发明的构思在现有技术的基础上通过逻辑分析、推理或者有限的实验可以得到的技术方案,皆应在由权利要求书所确定的保护范围内。

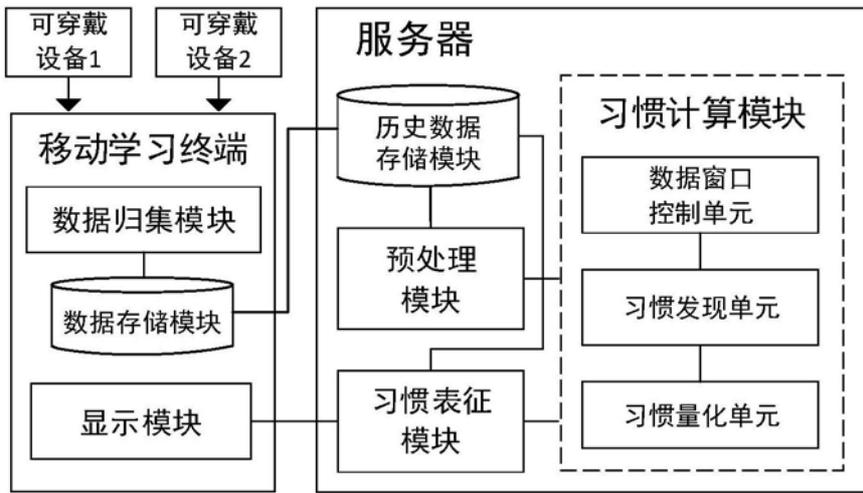


图1

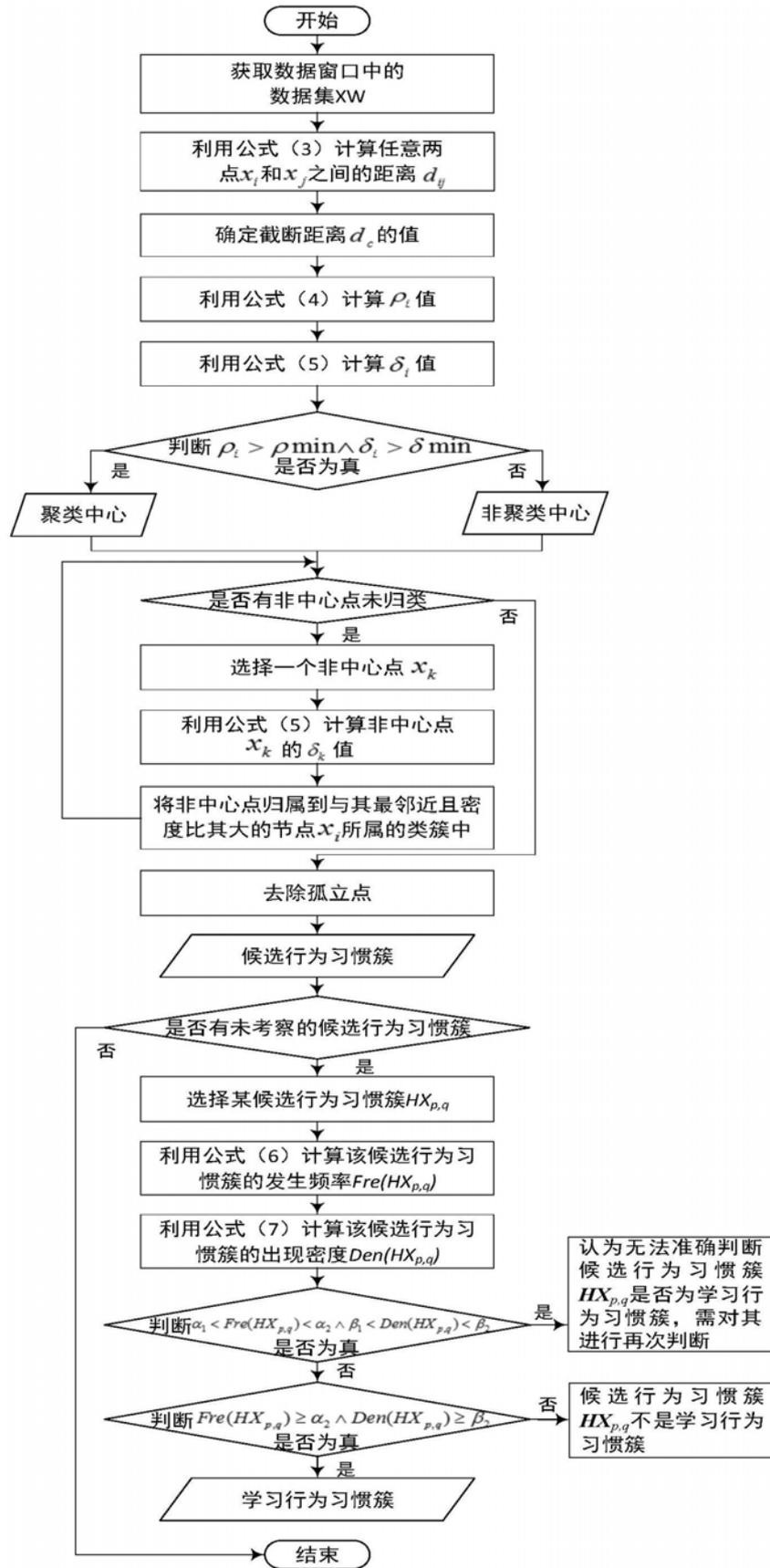


图2

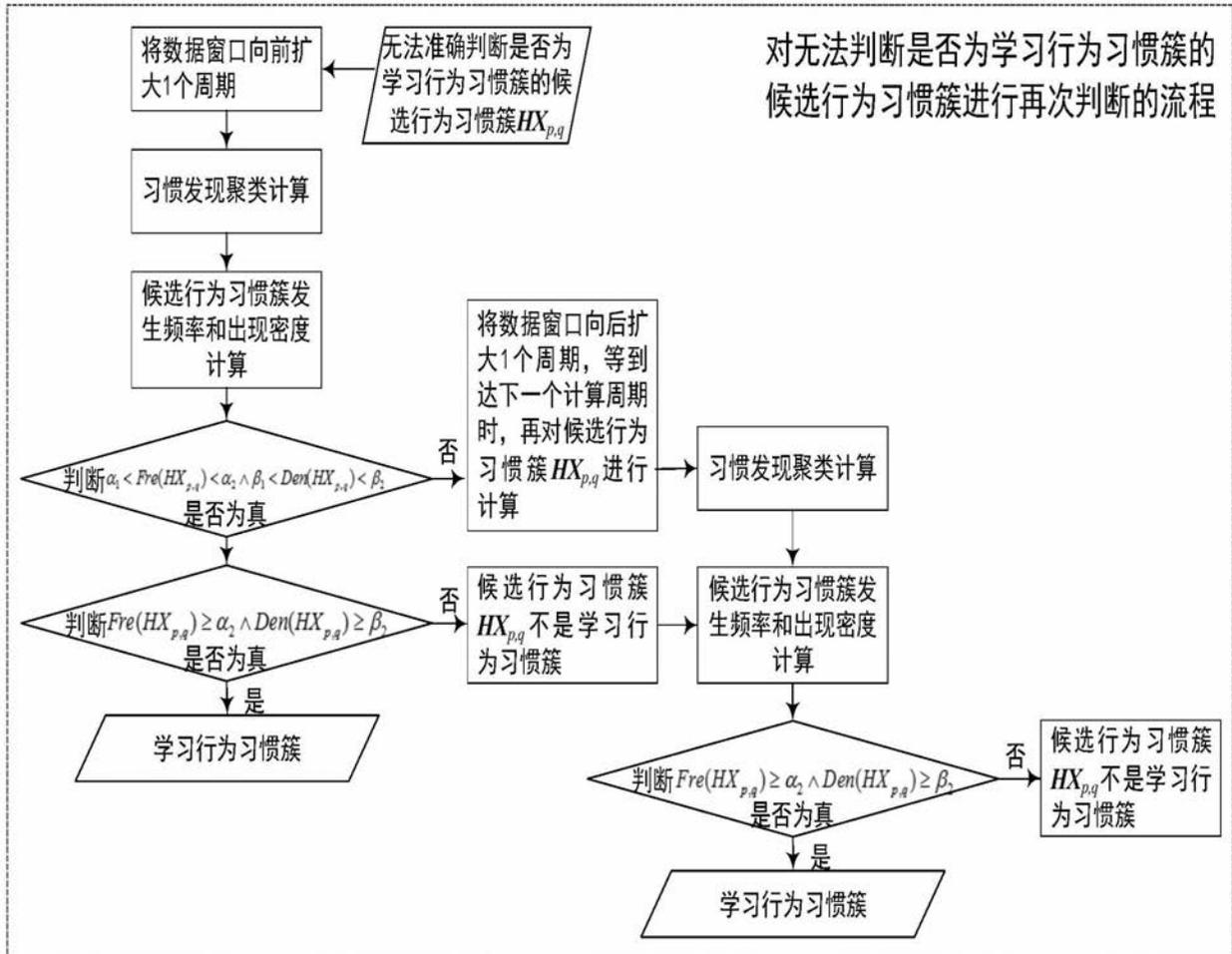


图3