



(12) 发明专利

(10) 授权公告号 CN 109426732 B

(45) 授权公告日 2021.09.21

(21) 申请号 201710721540.2

G06F 21/60 (2013.01)

(22) 申请日 2017.08.22

(56) 对比文件

(65) 同一申请的已公布的文献号
申请公布号 CN 109426732 A

CN 102271124 A, 2011.12.07
CN 102271124 A, 2011.12.07
CN 106445936 A, 2017.02.22

(43) 申请公布日 2019.03.05

CN 106237617 A, 2016.12.21
CN 106027257 A, 2016.10.12

(73) 专利权人 创新先进技术有限公司
地址 开曼群岛大开曼岛西湾路802号木槿
街大展览馆31119号邮箱邮编KY1-
1205

CN 106997439 A, 2017.08.01
US 2010235635 A1, 2010.09.16

审查员 张莹

(72) 发明人 解岭 李小龙

(74) 专利代理机构 北京晋德允升知识产权代理
有限公司 11623

代理人 王戈

(51) Int. Cl.

G06F 21/62 (2013.01)

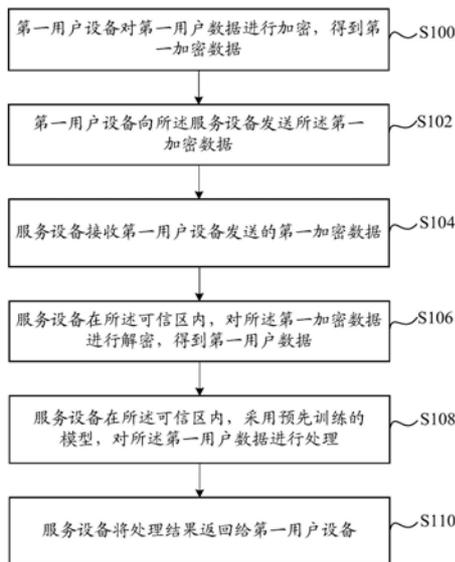
权利要求书4页 说明书12页 附图5页

(54) 发明名称

一种数据处理方法及装置

(57) 摘要

本说明书实施例公开了一种数据处理方法及装置。在本说明书实施例中,在服务设备的可信区内对用户设备发送的加密数据进行解密得到用户数据,以及在所述可信区内采用预先训练的模型对用户数据进行处理,并将处理结果返回给用户设备。



1. 一种数据处理方法, 服务设备中具有可信区, 所述可信区位于所述服务设备的CPU中, 所述服务设备中包括多个第一CPU; 所述服务设备还包括至少一个第二CPU; 所述方法包括:

所述服务设备接收第一用户设备发送的第一加密数据;

在所述可信区内, 对所述第一加密数据进行解密, 得到第一用户数据;

在所述可信区内, 采用预先训练的模型对所述第一用户数据进行处理;

将处理结果返回给所述第一用户设备;

所述在所述可信区内, 采用预先训练的模型对所述第一用户数据进行处理, 具体包括:

针对每个第二用户设备, 根据所述第一用户数据对应的数据标识, 从该第二用户设备中获取第二加密数据, 其中, 所述第二加密数据是该第二用户设备对所述数据标识对应的第二用户数据进行加密得到的;

在所述可信区内, 对获取的第二加密数据进行解密, 得到第二用户数据;

在所述可信区内, 采用预先训练的模型对所述第一用户数据和所述第二用户数据进行处理;

所述采用预先训练的模型对所述第一用户数据和所述第二用户数据进行处理, 具体包括:

分别在不同的第一CPU的可信区内, 采用预先训练的模型对所述第一用户数据和各第二用户数据进行处理, 得到所述第一用户数据和各第二用户数据分别对应的中间结果, 通过所述第二CPU, 根据各第一CPU分别发送的中间结果, 确定处理结果。

2. 根据权利要求1所述的方法, 在所述可信区内, 对所述第一加密数据进行解密, 得到第一用户数据, 具体包括:

在第一CPU的可信区内, 对所述第一加密数据进行解密, 得到第一用户数据;

在所述可信区内, 对获取的第二加密数据进行解密, 得到第二用户数据, 具体包括:

分别在不同的第一CPU的可信区内, 对不同的各第二加密数据进行解密, 得到第二用户数据; 用于解密第二加密数据的第一CPU不同于用于解密第一加密数据的第一CPU。

3. 根据权利要求1所述的方法, 在所述可信区内, 采用预先训练的模型对所述第一用户数据进行处理, 具体包括:

在所述可信区内, 采用预先训练的模型对所述第一用户数据进行处理, 得到所述第一用户数据对应的中间结果;

根据所述第一用户数据对应的数据标识, 在预先存储的中间结果中, 选择所述数据标识对应的中间结果;

根据所述第一用户数据对应的中间结果和从预先存储的中间结果中选择的中间结果, 确定处理结果。

4. 根据权利要求3所述的方法, 预先存储中间结果, 具体包括:

预先针对每个第二用户设备, 从该第二用户设备中获取第二加密数据, 其中, 所述第二加密数据是该第二用户设备对各数据标识分别对应的第二用户数据进行加密得到的;

在所述可信区内, 对获取的第二加密数据进行解密, 得到各数据标识分别对应的第二用户数据;

在所述可信区内, 采用预先训练的模型对每个第二用户数据进行处理, 得到每个第二

用户数据分别对应的中间结果；

存储每个第二用户数据的数据标识和对应的中间结果。

5. 一种数据处理方法, 服务设备中具有可信区, 所述可信区位于所述服务设备的CPU中, 所述服务设备中包括多个第一CPU; 所述服务设备还包括至少一个第二CPU; 所述方法包括:

用户设备对用户数据进行加密, 得到加密数据;

向所述服务设备发送所述加密数据, 以使所述服务设备在所述可信区内, 对所述加密数据进行解密, 得到所述用户数据, 并采用预先训练的模型对所述用户数据进行处理;

接收所述服务设备返回的处理结果;

所述采用预先训练的模型对所述用户数据进行处理, 具体包括:

针对每个用户设备, 根据所述用户数据对应的数据标识, 从该用户设备中获取加密数据, 其中, 所述加密数据是该用户设备对所述数据标识对应的用户数据进行加密得到的;

在所述可信区内, 对获取的加密数据进行解密, 得到用户数据;

在所述可信区内, 采用预先训练的模型对所述用户数据进行处理;

所述在所述可信区内, 采用预先训练的模型对所述用户数据进行处理, 具体包括:

分别在不同的第一CPU的可信区内, 采用预先训练的模型对所述用户数据进行处理, 得到所述用户数据对应的中间结果, 通过所述第二CPU, 根据所述第一CPU发送的中间结果, 确定处理结果。

6. 一种数据处理装置, 所述装置中具有可信区, 所述可信区位于所述装置的CPU中, 所述装置中包括多个第一CPU; 所述装置还包括至少一个第二CPU; 所述装置包括:

接收模块, 接收第一用户设备发送的第一加密数据;

解密模块, 在所述可信区内, 对所述第一加密数据进行解密, 得到第一用户数据;

处理模块, 在所述可信区内, 采用预先训练的模型对所述第一用户数据进行处理;

返回模块, 将处理结果返回给所述第一用户设备;

所述处理模块, 针对每个第二用户设备, 根据所述第一用户数据对应的数据标识, 从该第二用户设备中获取第二加密数据, 其中, 所述第二加密数据是该第二用户设备对所述数据标识对应的第二用户数据进行加密得到的; 在所述可信区内, 对获取的第二加密数据进行解密, 得到第二用户数据; 在所述可信区内, 采用预先训练的模型对所述第一用户数据和所述第二用户数据进行处理;

所述处理模块, 分别在不同的第一CPU的可信区内, 采用预先训练的模型对所述第一用户数据和各第二用户数据进行处理, 得到所述第一用户数据和各第二用户数据分别对应的中间结果, 通过所述第二CPU, 根据各第一CPU分别发送的中间结果, 确定处理结果。

7. 根据权利要求6所述的装置, 所述处理模块, 采用预先训练的模型分别对所述第一用户数据和第二用户数据进行处理, 得到所述第一用户数据和各第二用户数据分别对应的中间结果; 根据各中间结果, 确定处理结果。

8. 根据权利要求7所述的装置,

所述解密模块, 在第一CPU的可信区内, 对所述第一加密数据进行解密, 得到第一用户数据;

所述处理模块, 分别在不同的第一CPU的可信区内, 对不同的各第二加密数据进行解

密,得到第二用户数据;用于解密第二加密数据的第一CPU不同于用于解密第一加密数据的第一CPU。

9. 根据权利要求6所述的装置,所述处理模块,在所述可信区内,采用预先训练的模型对所述第一用户数据进行处理,得到所述第一用户数据对应的中间结果;根据所述第一用户数据对应的数据标识,在预先存储的中间结果中,选择所述数据标识对应的中间结果;根据所述第一用户数据对应的中间结果和从预先存储的中间结果中选择的中间结果,确定处理结果。

10. 根据权利要求9所述的装置,所述装置还包括:

预处理模块,预先针对每个第二用户设备,从该第二用户设备中获取第二加密数据,其中,所述第二加密数据是该第二用户设备对各数据标识分别对应的第二用户数据进行加密得到的;在所述可信区内,对获取的第二加密数据进行解密,得到各数据标识分别对应的第二用户数据;在所述可信区内,采用预先训练的模型对每个第二用户数据进行处理,得到每个第二用户数据分别对应的中间结果;存储每个第二用户数据的数据标识和对应的中间结果。

11. 一种数据处理装置,服务设备中具有可信区,所述可信区位于所述服务设备的CPU中,所述服务设备中包括多个第一CPU;所述服务设备还包括至少一个第二CPU;所述装置包括:

加密模块,对用户数据进行加密,得到加密数据;

发送模块,向所述服务设备发送所述加密数据,以使所述服务设备在所述可信区内,对所述加密数据进行解密,得到所述用户数据,并采用预先训练的模型对所述用户数据进行处理;

接收模块,接收所述服务设备返回的处理结果;

所述发送模块,针对每个用户设备,根据所述用户数据对应的数据标识,从该用户设备中获取加密数据,其中,所述加密数据是该用户设备对所述数据标识对应的用户数据进行加密得到的;

在所述可信区内,对获取的加密数据进行解密,得到用户数据;

在所述可信区内,采用预先训练的模型对所述用户数据进行处理;

所述发送模块,分别在不同的第一CPU的可信区内,采用预先训练的模型对所述用户数据进行处理,得到所述用户数据对应的中间结果,通过所述第二CPU,根据所述第一CPU发送的中间结果,确定处理结果。

12. 一种服务设备,所述服务设备中具有可信区,所述可信区位于所述服务设备的CPU中,所述服务设备中包括多个第一CPU;所述服务设备还包括至少一个第二CPU;所述服务设备包括一个或多个处理器及存储器,所述存储器存储有程序,并且被配置成由所述一个或多个处理器执行以下步骤:

接收第一用户设备发送的第一加密数据;

在所述可信区内,对所述第一加密数据进行解密,得到第一用户数据;

在所述可信区内,采用预先训练的模型对所述第一用户数据进行处理;

将处理结果返回给所述第一用户设备;

所述在所述可信区内,采用预先训练的模型对所述第一用户数据进行处理,具体包括:

针对每个第二用户设备,根据所述第一用户数据对应的数据标识,从该第二用户设备中获取第二加密数据,其中,所述第二加密数据是该第二用户设备对所述数据标识对应的第二用户数据进行加密得到的;

在所述可信区内,对获取的第二加密数据进行解密,得到第二用户数据;

在所述可信区内,采用预先训练的模型对所述第一用户数据和所述第二用户数据进行处理;

所述采用预先训练的模型对所述第一用户数据和所述第二用户数据进行处理,具体包括:

分别在不同的第一CPU的可信区内,采用预先训练的模型对所述第一用户数据和各第二用户数据进行处理,得到所述第一用户数据和各第二用户数据分别对应的中间结果,通过所述第二CPU,根据各第一CPU分别发送的中间结果,确定处理结果。

13. 一种用户设备,服务设备中具有可信区,所述可信区位于所述服务设备的CPU中,所述服务设备中包括多个第一CPU;所述服务设备还包括至少一个第二CPU;所述用户设备包括一个或多个处理器及存储器,所述存储器存储有程序,并且被配置成由所述一个或多个处理器执行以下步骤:

对用户数据进行加密,得到加密数据;

向所述服务设备发送所述加密数据,以使所述服务设备在所述可信区内,对所述加密数据进行解密,得到所述用户数据,并采用预先训练的模型对所述用户数据进行处理;

接收所述服务设备返回的处理结果;

所述采用预先训练的模型对所述用户数据进行处理,具体包括:

针对每个用户设备,根据所述用户数据对应的数据标识,从该用户设备中获取加密数据,其中,所述加密数据是该用户设备对所述数据标识对应的用户数据进行加密得到的;

在所述可信区内,对获取的加密数据进行解密,得到用户数据;

在所述可信区内,采用预先训练的模型对所述用户数据进行处理;

所述在所述可信区内,采用预先训练的模型对所述用户数据进行处理,具体包括:

分别在不同的第一CPU的可信区内,采用预先训练的模型对所述用户数据进行处理,得到所述用户数据对应的中间结果,通过所述第二CPU,根据所述第一CPU发送的中间结果,确定处理结果。

一种数据处理方法及装置

技术领域

[0001] 本说明书涉及数据挖掘技术领域,尤其涉及一种数据处理方法及装置。

背景技术

[0002] 目前,服务提供商使用自己训练的模型为企业或个人等用户提供预测、风控、预警等服务的模式已经日渐成熟。在这种模式下,服务提供商获取用户提供的用户数据,采用预先训练的模型对获取的用户数据进行处理,进而为用户提供服务。但是,用户数据往往涉及用户的隐私,用户通常并不想将自己的隐私泄露出去。

[0003] 基于现有技术,需要一种数据处理方法,在保护用户的隐私不泄露的前提下,使用模型对用户数据进行处理。

发明内容

[0004] 本说明书实施例提供一种数据处理方法及装置,以解决如何在保护用户的隐私不泄露的前提下,使用模型对用户数据进行处理的问题。

[0005] 为解决上述技术问题,本说明书实施例是这样实现的:

[0006] 本说明书实施例提供的一种业务执行方法,服务设备中具有可信区,所述方法包括:

[0007] 所述服务设备接收第一用户设备发送的第一加密数据;

[0008] 在所述可信区内,对所述第一加密数据进行解密,得到第一用户数据;

[0009] 在所述可信区内,采用预先训练的模型对所述第一用户数据进行处理;

[0010] 将处理结果返回给所述第一用户设备。

[0011] 本说明书实施例提供的另一种数据处理方法,服务设备中具有可信区,所述方法包括:

[0012] 用户设备对用户数据进行加密,得到加密数据;

[0013] 向所述服务设备发送所述加密数据,以使所述服务设备在所述可信区内,对所述加密数据进行解密,得到所述用户数据,并采用预先训练的模型对所述用户数据进行处理;

[0014] 接收所述服务设备返回的处理结果。

[0015] 本说明书实施例提供的一种数据处理装置,所述装置中具有可信区,所述装置包括:

[0016] 接收模块,接收第一用户设备发送的第一加密数据;

[0017] 解密模块,在所述可信区内,对所述第一加密数据进行解密,得到第一用户数据;

[0018] 处理模块,在所述可信区内,采用预先训练的模型对所述第一用户数据进行处理;

[0019] 返回模块,将处理结果返回给所述第一用户设备。

[0020] 本说明书实施例提供的另一种数据处理装置,服务设备中具有可信区,所述装置包括:

[0021] 加密模块,对用户数据进行加密,得到加密数据;

[0022] 发送模块,向所述服务设备发送所述加密数据,以使所述服务设备在所述可信区内,对所述加密数据进行解密,得到所述用户数据,并采用预先训练的模型对所述用户数据进行处理;

[0023] 接收模块,接收所述服务设备返回的处理结果。

[0024] 本说明书实施例提供的一种服务设备,所述服务设备中具有可信区,所述服务设备包括一个或多个处理器及存储器,所述存储器存储有程序,并且被配置成由所述一个或多个处理器执行以下步骤:

[0025] 接收第一用户设备发送的第一加密数据;

[0026] 在所述可信区内,对所述第一加密数据进行解密,得到第一用户数据;

[0027] 在所述可信区内,采用预先训练的模型对所述第一用户数据进行处理;

[0028] 将处理结果返回给所述第一用户设备。

[0029] 本说明书实施例提供的一种用户设备,服务设备中具有可信区,所述用户设备包括一个或多个处理器及存储器,所述存储器存储有程序,并且被配置成由所述一个或多个处理器执行以下步骤:

[0030] 对用户数据进行加密,得到加密数据;

[0031] 向所述服务设备发送所述加密数据,以使所述服务设备在所述可信区内,对所述加密数据进行解密,得到所述用户数据,并采用预先训练的模型对所述用户数据进行处理;

[0032] 接收所述服务设备返回的处理结果。

[0033] 由以上本说明书实施例提供的技术方案可见,在本说明书实施例中,在服务设备的可信区内对用户设备发送的加密数据进行解密得到用户数据,以及在所述可信区内采用预先训练的模型对用户数据进行处理。如此一来,用户数据只暴露在可信区内,即便是服务设备的拥有者,也无法获取可信区内的用户数据,用户隐私也就不会泄露。

附图说明

[0034] 为了更清楚地说明本说明书实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本说明中记载的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动性的前提下,还可以根据这些附图获得其他的附图。

[0035] 图1是本说明书实施例提供的一种数据处理方法流程图;

[0036] 图2是本说明书实施例提供的数据处理系统架构图;

[0037] 图3是本说明书实施例提供的多个用户设备存储的用户数据示意图;

[0038] 图4是本说明书实施例提供的一种数据处理方法示意图;

[0039] 图5是本说明书实施例提供的另一种数据处理方法示意图;

[0040] 图6是本说明书实施例提供的一种数据处理装置示意图;

[0041] 图7是本说明书实施例提供的另一种数据处理装置示意图;

[0042] 图8是本说明书实施例提供的服务设备示意图;

[0043] 图9是本说明书实施例提供的用户设备示意图。

具体实施方式

[0044] 目前,服务提供商一般采用预先训练的模型对用户提供的用户数据进行处理,将处理结果返回给用户。通过这种方式,服务提供商可以对用户提供预测、风控、预警等服务。例如,企业用户A将自己的财务报表提供给服务提供商,请求服务提供商预测其下个财年的盈亏情况。服务提供商采用预先训练的用于预测企业盈亏情况的模型,对企业用户A提供的用户数据进行处理,得到企业用户A下个财年的净利润的期望值,作为处理结果返回给企业用户A。

[0045] 但是,用户数据往往涉及用户的隐私(在上例中,企业用户A的财务报表中显然会涉及企业用户A的隐私)。

[0046] 为了防止用户的隐私泄露,在本说明书的一个或多个实施例中,将服务设备中具有的可信区作为与外部隔离的执行环境,在所述可信区中对加密数据进行解密得到用户数据,以及在可信区内采用预先训练的模型对用户数据进行处理,最终由服务设备输出处理结果,使得在整个数据处理过程中,用户数据始终不会暴露给可信区之外,从而保护了用户的隐私。其中,所述可信区可以具体位于服务设备的中央处理器(Central Processing Unit,CPU)内,也可以具体位于服务设备的其他组件内,例如图形处理器(Graphics Processing Unit,GPU)、现场可编程门阵列(Field-Programmable Gate Array,FPGA)等芯片中。为了描述的方便,在本说明书的一个或多个实施例中,以服务设备中的可信区具体位于服务设备的CPU中为例展开说明,但本领域技术人员应当理解,这并不构成对本说明书所要求保护的技术方案的限制。

[0047] 为了使本技术领域的人员更好地理解本说明书中的技术方案,下面将结合本说明书一个或多个实施例中的附图,对本说明书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说明书一部分实施例,而不是全部的实施例。基于本说明书实施例,本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例,都应当属于本说明书保护的范围。

[0048] 以下结合附图,详细说明本说明书各实施例提供的技术方案。

[0049] 图1是本说明书实施例提供的业务执行方法流程图,包括以下步骤:

[0050] S100:第一用户设备对第一用户数据进行加密,得到第一加密数据。

[0051] S102:第一用户设备向所述服务设备发送所述第一加密数据。

[0052] 在本说明书实施例中,所述第一用户设备是向服务设备主动发送加密数据的用户设备,也就是对基于模型的预测、风控、预警等服务有需求的用户的设备。为了描述的方便,将第一用户数据发送的加密数据称为第一加密数据,第一加密数据是第一用户设备对自身存储的第一用户数据加密得到的。下文的第二用户数据是和第一用户数据具有关联性(对应于同一数据标识)的用户数据,下文的第二用户设备是存储有第二用户数据的用户设备。

[0053] 服务设备即是服务提供商的设备,是模型的拥有者。服务设备可采用预先训练的模型对第一用户设备存储的第一用户数据进行处理,以对用户提供预测、风控、预警等服务。

[0054] 例如,某个企业用户想要预测自己未来的盈亏情况,服务提供商拥有预测企业盈亏情况的模型,可以提供相应的预测服务,那么该企业用户的服务器即为第一用户设备,服务提供商的服务器即为服务设备。

[0055] 在本说明书实施例中,当第一用户设备对服务提供商提供的服务有需求时,为了防止用户的隐私泄露,第一用户设备先对其存储的第一用户数据进行加密,得到第一加密数据,再向服务设备发送所述第一加密数据。

[0056] S104:服务设备接收第一用户设备发送的第一加密数据。

[0057] S106:服务设备在所述可信区内,对所述第一加密数据进行解密,得到第一用户数据。

[0058] S108:服务设备在所述可信区内,采用预先训练的模型,对所述第一用户数据进行处理。

[0059] 在本说明书实施例中,服务设备中具有可信区。服务设备在接收到第一用户设备发送的第一加密数据之后,可以在可信区内对接收到的第一加密数据进行解密得到用户数据,并继续在可信区内对解密得到的第一用户数据进行处理。在图1所示的各步骤中,用户数据仅会暴露在可信区之内。

[0060] 为本领域技术人员所熟知的是,可信区可以是在CPU的存储区域中物理隔离出的区域,并且在所述可信区内可以执行指定操作。在本说明书实施例中,在可信区内执行的指定操作包括对加密数据进行解密的解密操作和对解密得到的用户数据进行处理的处理操作,而所述指定操作不包括从可信区内提取用户数据等容易造成用户数据从可信区中泄露出去的操作。通过所述可信区技术,服务设备可以在不泄露用户数据的前提下,对用户数据进行处理,进而为用户提供预测、风控等服务。

[0061] 需要说明的是,在本说明书实施例中,可以采用英特尔公司开发的可信区技术(Intel Software Guard Extensions),通常简称为“SGX”。当然,其他类似的可信区技术也可以应用于本方案,不再赘述。

[0062] 图2是本说明书实施例提供的数据处理系统架构图。如图2所示的架构,当采用SGX技术时,第一用户设备向服务设备发送第一加密数据、可信区定义文件(.ed1文件)以及动态链接库文件(.dll文件或.so文件),服务设备将接收到的第一加密数据存入存储器,以待后续处理。同时,服务设备中的加载器根据所述可信区定义文件在CPU中定义可信区,并根据动态链接库文件,在定义的可信区内执行上述指定操作(解密操作和处理操作)。

[0063] 在本说明书实施例中,上述处理操作可以是在可信区内将解密得到的第一用户数据作为预先训练的模型的输入,采用所述模型对第一用户数据进行处理。具体而言,所述模型的模型参数可以是用户数据中每条数据的权重,而处理结果则是对各条数据计算加权和。例如,模型参数为 $W = (w_1, w_2, \dots, w_n)$,这 n 个权重值 $w_1 \sim w_n$ 一一对应于解密得到的用户

设备A提供的用户数据中的 n 条数据 $a_1 \sim a_n$ 。那么,处理结果即是 $\text{sum}A = \sum_{i=1}^n w_i a_i$ 。

[0064] S110:服务设备将处理结果返回给第一用户设备。

[0065] 在本说明书实施例中,服务设备在可信区内执行了上述处理操作之后,将处理结果返回给所述第一用户设备。

[0066] 通过图1所示的数据处理方法,在服务设备的可信区内执行指定操作:1、对加密数据进行解密得到用户数据;2、对解密得到的用户数据进行处理。由于执行的指定操作不包括从可信区内提取用户数据等容易造成用户数据从可信区泄露出去的操作,因此,用户数据只暴露在可信区内,即便是服务设备的拥有者,也无法获取可信区内的用户数据,用户的

隐私也就不会泄露。此外,由于对用户数据的处理操作在所述可信区内执行,用于处理用户数据的模型也不会泄露或被修改。

[0067] 此外,在本说明书的一个或多个实施例中,为了防止有人恶意操纵服务设备在所述可信区之外执行提取用户数据的操作,在步骤S100中,用户设备可以采用在所述用户设备的CPU的可信区内执行的加密算法,对用户数据进行加密。

[0068] 事实上,由于CPU的可信区实际上是在CPU的存储区域中物理隔离出的区域,因此可信区具有特定的物理参数。而在CPU的可信区内执行的加密算法,即是需要使用可信区的物理参数的加密算法。相应地,在步骤S106中,服务设备可以在其CPU的可信区内采用与所述加密算法对应的解密算法对接收到的加密数据进行解密。与所述加密算法对应的解密算法应在所述服务设备的CPU的可信区内执行。

[0069] 显然,解密算法中需要使用的可信区的物理参数与所述加密算法中使用的可信区的物理参数应当一致。也就是说,所述用户设备的CPU与所述服务设备的CPU的可信区的物理参数应当一致。一般的,同一CPU制造商制造的CPU可满足此要求。

[0070] 另外,通过图1所示的方法得到的处理结果可以是对企业的业务的预测结果(如盈利或亏损,又如净利润的期望值)、风险评价(如风险概率),也可以是对个人的信用评价(如信用评级、信用打分)。

[0071] 以个人信用评分的业务场景为例,倘若银行用户A需要获知某人甲的信用评分,以决定是否向甲提供贷款,那么,银行用户A的服务器(第一用户设备)将自己存储的用户数据(可以是甲在银行用户A处产生的信用记录)加密后提供给服务设备,服务设备在可信区内采用预先训练的信用评分模型对解密得到的用户数据进行处理,计算甲的信用评分。随后,服务设备可以将甲的信用评分,作为处理结果返回给银行用户A的服务器。

[0072] 进一步地,继续以个人信用评分的业务场景为例,某人往往不止在一个机构产生信用记录(某人在某个机构产生的信用记录即是该机构的服务器存储的用户数据)。在上述举例中,甲除了在银行用户A处产生了信用记录,也可能在其他机构(其他银行、金融公司等)也产生了信用记录。众所周知,针对个人而言,服务设备获取的这个人的信用记录越完整,对这个人的信用记录进行处理后,得到的这个人的信用评分就越准确。在上述举例中,银行用户A将自己存储的甲的信用记录提供给服务提供商,请求服务提供商对甲进行信用评分,为了信用评分尽可能准确,服务提供商可以进一步向其他机构获取甲产生的信用记录,综合多个机构提供的甲的信用记录,来对甲进行信用评分。

[0073] 实际上,不止限于个人信用评分的业务场景,凡是涉及采用模型对个人进行行为预测、评价的业务场景,都符合上述情况。在这些业务场景下,某个人往往在多个机构(用户)都产生了数据记录,这多个机构的用户设备分别存储的用户数据可组成这个人的更为完整的数据记录。

[0074] 图3是本说明书实施例提供的多个用户设备存储的用户数据示意图。如图3所示,用户设备A、B、C分别是机构A、B、C的服务器,以用户设备A为例,用户设备A存储的用户数据即为个人1~k在机构A产生的数据记录,其中,第一行用户数据为个人1的数据记录,第2行用户数据为个人2的数据记录,以此类推。同样地,用户设备B、用户设备C也分别存储有个人1~k的数据记录。

[0075] 可见,对于某个人而言,这个人在机构A~C处都产生了数据记录。倘若机构A(第一

用户设备)请求服务提供商对个人1的行为进行预测,那么服务提供商除了对机构A提供的个人1的数据记录进行处理之外,还可以进一步获取并处理机构B和机构C(第二用户设备)存储的个人1的数据记录(同一个人在不同的机构产生的数据记录具有关联性),以使最终得到的预测结果(也就是处理结果)尽可能准确。显然,第二用户设备存储的用户数据也会涉及隐私,因此,服务设备从第二用户设备获取的通常也是加密数据。

[0076] 针对上述的业务场景,为了在各用户设备存储的用户数据中的隐私不泄露的前提下,采用模型对各用户设备提供的用户数据进行处理,本说明书实施例提供了以下两种方式。值得强调的是,除以下两种方式之外,其他基于可信区技术,在可信区内对多个用户设备发送的加密数据进行解密(各用户设备发送的各加密数据分别是对具有关联性的各用户数据进行加密得到的),以及在可信区内对解密得到的多个具有关联性的用户数据进行处理的方式,都在本技术方案所要求的保护范围之内。

[0077] 方式一

[0078] 服务设备可以在所述可信区内,采用预先训练的模型对来自多个用户设备的用户数据进行处理。这多个用户设备中不仅包括第一用户设备,也包括第二用户设备。如前所述,第一用户设备主动向服务设备发送第一加密数据,以请求服务设备采用模型处理对第一加密数据进行解密得到的第一用户数据,并对第一用户数据进行处理,而第二用户设备中存储有与第一用户数据有关联的第二用户数据(例如,同一个人在不同用户设备中产生的用户数据具有关联性),服务提供者为了更好地为第一用户提供更好的服务,主动向第二用户设备请求获取第二加密数据。

[0079] 具体而言,在步骤S106中,服务设备在可信区内对第一用户设备发送的加密数据进行解密得到第一用户数据;在步骤S108中,服务设备针对每个第二用户设备,根据所述第一用户数据对应的数据标识,从该第二用户设备中获取第二加密数据,其中,所述第二加密数据是该第二用户设备对所述数据标识对应的第二用户数据进行加密得到的;然后,服务设备在所述可信区内,对获取的第二加密数据进行解密,得到第二用户数据;进而在所述可信区内,采用预先训练的模型对所述第一用户数据和所述第二用户数据进行处理。

[0080] 其中,同一数据标识在不同的用户设备中对应的用户数据具有上述关联性。例如,在个人信用评分的业务场景下,张三在机构A、B、C产生的数据记录都对应于同一数据标识,该数据标识可以是张三的手机号、身份证号等。需要说明的是,第一用户设备向服务设备发送第一加密数据时,可同时将第一用户数据对应的数据标识也发送给服务设备。当然,第一加密数据也可以是第一用户设备对第一用户数据和第一用户数据的数据标识进行加密得到的,这样一来,服务设备在可信区内解密第一加密数据后,就获得了第一用户数据的数据标识。

[0081] 进一步地,服务设备可以在所述可信区内,采用预先训练的模型分别对所述第一用户数据和第二用户数据进行处理,得到所述第一用户数据和各第二用户数据分别对应的中间结果,再根据各中间结果,确定处理结果。

[0082] 图4是本说明书实施例提供的一种数据处理方法示意图。如图4所示,第一用户设备是主动向服务设备发送加密数据的一方,第一用户设备请求服务设备对个人1的行为进行预测,那么第一用户设备发送给服务设备的第一加密数据可以是对个人1在第一用户设备对应的用户处产生的数据记录(用户数据)加密得到的。此外,第二用户设备中也存储有

个人1产生的数据记录。

[0083] 继续参见图4,服务设备中包括至少一个第一CPU、至少一个第二CPU。服务设备在第一CPU的可信区内,对所述第一加密数据进行解密,得到第一用户数据,以及,分别在不同的第一CPU的可信区内,对不同的各第二加密数据进行解密,得到第二用户数据。其中,用于解密第二加密数据的第一CPU不同于用于解密第一加密数据的第一CPU。

[0084] 然后,服务设备可以分别在不同的第一CPU的可信区内,采用预先训练的模型对所述第一用户数据和各第二用户数据进行处理,得到所述第一用户数据和各第二用户数据分别对应的中间结果,并将得到的各中间结果发送给所述第二CPU。进而通过所述第二CPU,根据各第一CPU分别发送的中间结果,确定处理结果。

[0085] 具体而言,如图4所示,在对个人进行信用评价的业务场景下,评价一个人的信用需要 n 个维度的数据 $x_{11} \sim x_{1n}$, x_{11} 可以是有车产, x_{12} 可以是有房产, \dots x_{1n} 可以是有大学学历。个人1产生的第1个维度 \sim 第 t 个维度的数据,作为第一用户数据存储在第一用户设备中,个人1产生的第 $t+1$ 个维度 \sim 第 n 个维度的数据,作为第二用户数据存储在第二用户设备中。

[0086] 而所述模型的模型参数可以是每条数据对应的权重。由于第一用户数据中包括 $x_{11} \sim x_{1t}$,因此,在左边的第一CPU中,可以仅使用相应的模型参数 $w_1 \sim w_t$ 计算 sum1 即可。当然,在左边的第一CPU中,也可以使用模型参数 $w_1 \sim w_n$ 计算 sum1 ,由于第一用户数据中,第 $t+1$ 个维度 \sim 第 n 个维度的数据为空,可视为取0,因此,对计算得到的 sum1 也不会产生影响。

[0087] 继续参见图4,服务设备在左边的第一CPU中根据第一用户数据和模型参数计算得到第1个维度 \sim 第 t 个维度的数据的加权和(sum1),在右边的第一CPU中根据第二用户数据和模型参数计算得到第 $t+1$ 个维度 \sim 第 n 个维度的数据的加权和(sum2),再由第二CPU综合 sum1 和 sum2 ,得到最终的处理结果 $f(\text{sum1}, \text{sum2})$,返回给第一用户设备。

[0088] 方式二

[0089] 图5是本说明书实施例提供的另一种数据处理方法示意图。如图5所示,服务设备可以预先针对每个第二用户设备,从该第二用户设备中获取第二加密数据,其中,所述第二加密数据是该第二用户设备对各数据标识分别对应的第二用户数据进行加密得到的。然后,在所述可信区内,对获取的第二加密数据进行解密,得到各数据标识分别对应的第二用户数据,接着在所述可信区内,采用预先训练的模型对每个第二用户数据进行处理,得到每个第二用户数据分别对应的中间结果,最后存储每个第二用户数据的数据标识和对应的中间结果。

[0090] 需要说明的是,上述预先执行的操作,可以是在服务设备中同一个CPU的可信区内,依次针对各第二加密数据执行的,也可以是分别在服务设备中不同的CPU的可信区内针对不同的第二加密数据执行的。

[0091] 当服务设备接收到第一用户设备发送的第一加密数据时,可以在可信区内针对第一加密数据执行指定操作,即对第一解密数据进行解密得到第一用户数据,以及对采用预先训练的模型对第一用户数据进行处理,得到第一用户设备对应的中间结果。然后,根据所述第一用户数据对应的数据标识,在存储的中间结果中,确定所述数据标识对应的中间结果,也即在预先存储的各第二用户设备分别对应的中间结果中,选择与第一用户数据关联的第二用户数据对应的中间结果。最后,根据所述第一用户数据对应的中间结果和确定的

中间结果,确定处理结果,并将处理结果返回给第一用户设备。

[0092] 在方式二中,也可采用图4中使用的计算方法,预先针对解密得到的每个第二用户数据,根据模型参数计算该第二用户数据中包含的各维度的数据的加权和,作为该第二用户数据对应的中间结果,并存储该第二用户数据的数据标识和对应的中间结果。在服务设备接收到第一加密数据后,可以根据模型参数计算解密得到第一用户数据中包含的各维度的数据的加权和,作为第一用户数据对应的中间结果,并综合第一用户数据对应的中间结果和存储的与第一用户数据关联的第二用户数据对应的中间结果,得到最终的处理结果,返回给第一用户设备。

[0093] 通过图5所示的方式,服务设备在接收到第一加密数据,并解密第一加密数据得到第一用户数据后,无需在线与各第二用户设备通讯,获取各第二加密数据,而是直接调用存储的与第一用户数据关联的第二用户数据对应的中间结果即可,提升了数据处理的效率。

[0094] 基于图1所示的数据处理方法,本说明书实施例还对应提供了一种数据处理装置,如图6所示,所述装置中具有可信区,所述装置包括:

[0095] 接收模块601,接收第一用户设备发送的第一加密数据;

[0096] 解密模块602,在所述可信区内,对所述第一加密数据进行解密,得到第一用户数据;

[0097] 处理模块603,在所述可信区内,采用预先训练的模型对所述第一用户数据进行处理;

[0098] 返回模块604,将处理结果返回给所述第一用户设备。

[0099] 所述处理模块603,针对每个第二用户设备,根据所述第一用户数据对应的数据标识,从该第二用户设备中获取第二加密数据,其中,所述第二加密数据是该第二用户设备对所述数据标识对应的第二用户数据进行加密得到的;在所述可信区内,对获取的第二加密数据进行解密,得到第二用户数据;在所述可信区内,采用预先训练的模型对所述第一用户数据和所述第二用户数据进行处理。

[0100] 所述处理模块603,采用预先训练的模型分别对所述第一用户数据和第二用户数据进行处理,得到所述第一用户数据和各第二用户数据分别对应的中间结果;根据各中间结果,确定处理结果。

[0101] 所述可信区位于所述服务设备的CPU中,所述服务设备中包括多个第一CPU;

[0102] 所述解密模块602,在第一CPU的可信区内,对所述第一加密数据进行解密,得到第一用户数据;

[0103] 所述处理模块603,分别在不同的第一CPU的可信区内,对不同的各第二加密数据进行解密,得到第二用户数据;用于解密第二加密数据的第一CPU不同于用于解密第一加密数据的第一CPU。

[0104] 所述服务设备还包括至少一个第二CPU;

[0105] 所述处理模块603,分别在不同的第一CPU的可信区内,采用预先训练的模型对所述第一用户数据和各第二用户数据进行处理,得到所述第一用户数据和各第二用户数据分别对应的中间结果,并将得到的各中间结果发送给所述第二CPU;通过所述第二CPU,根据各第一CPU分别发送的中间结果,确定处理结果。

[0106] 所述装置还包括:预处理模块605,预先针对每个第二用户设备,从该第二用户设

备中获取第二加密数据,其中,所述第二加密数据是该第二用户设备对各数据标识分别对应的第二用户数据进行加密得到的;在所述可信区内,对获取的第二加密数据进行解密,得到各数据标识分别对应的第二用户数据;在所述可信区内,采用预先训练的模型对每个第二用户数据进行处理,得到每个第二用户数据分别对应的中间结果;存储每个第二用户数据的数据标识和对应的中间结果。

[0107] 所述处理模块603,在所述可信区内,采用预先训练的模型对所述第一用户数据进行处理,得到所述第一用户数据对应的中间结果;根据所述第一用户数据对应的数据标识,在存储的中间结果中,确定所述数据标识对应的中间结果;根据所述第一用户数据对应的中间结果和确定的中间结果,确定处理结果。

[0108] 基于图1所示的数据处理方法,本说明书实施例还对应提供了另一种数据处理装置,如图7所示,包括:

[0109] 加密模块701,对用户数据进行加密,得到加密数据;

[0110] 发送模块702,向所述服务设备发送所述加密数据,以使所述服务设备在所述可信区内,对所述加密数据进行解密,得到所述用户数据,并采用预先训练的模型对所述用户数据进行处理;

[0111] 接收模块703,接收所述服务设备返回的处理结果。

[0112] 基于图1所示的数据处理方法,本说明书实施例还对应提供了一种服务设备,如图8所示,所述服务设备中具有可信区,所述服务设备包括一个或多个处理器及存储器,所述存储器存储有程序,并且被配置成由所述一个或多个处理器执行以下步骤:

[0113] 接收第一用户设备发送的第一加密数据;

[0114] 在所述可信区内,对所述第一加密数据进行解密,得到第一用户数据;

[0115] 在所述可信区内,采用预先训练的模型对所述第一用户数据进行处理;

[0116] 将处理结果返回给所述第一用户设备。

[0117] 基于图1所示的数据处理方法,本说明书实施例还对应提供了一种用户设备,如图9所示,服务设备中具有可信区,所述用户设备包括一个或多个处理器及存储器,所述存储器存储有程序,并且被配置成由所述一个或多个处理器执行以下步骤:

[0118] 对用户数据进行加密,得到加密数据;

[0119] 向所述服务设备发送所述加密数据,以使所述服务设备在所述可信区内,对所述加密数据进行解密,得到所述用户数据,并采用预先训练的模型对所述用户数据进行处理;

[0120] 接收所述服务设备返回的处理结果。

[0121] 本说明书中的各个实施例均采用递进的方式描述,各个实施例之间相同相似的部分互相参见即可,每个实施例重点说明的都是与其他实施例的不同之处。尤其,对于图8所示的服务设备和图9所示的用户设备而言,由于其基本相似于方法实施例,所以描述的比较简单,相关之处参见方法实施例的部分说明即可。

[0122] 在20世纪90年代,对于一个技术的改进可以很明显地区分是硬件上的改进(例如,对二极管、晶体管、开关等电路结构的改进)还是软件上的改进(对于方法流程的改进)。然而,随着技术的发展,当今的很多方法流程的改进已经可以视为硬件电路结构的直接改进。设计人员几乎都通过将改进的方法流程编程到硬件电路中来得到相应的硬件电路结构。因此,不能说一个方法流程的改进就不能用硬件实体模块来实现。例如,可编程逻辑器件

(Programmable Logic Device,PLD) (例如现场可编程门阵列(Field Programmable Gate Array,FPGA))就是这样一种集成电路,其逻辑功能由用户对器件编程来确定。由设计人员自行编程来把一个数字系统“集成”在一片PLD上,而不需要请芯片制造厂商来设计和制作专用的集成电路芯片。而且,如今,取代手工地制作集成电路芯片,这种编程也多半改用“逻辑编译器(logic compiler)”软件来实现,它与程序开发撰写时所用的软件编译器相类似,而要编译之前的原始代码也得用特定的编程语言来撰写,此称之为硬件描述语言(Hardware Description Language,HDL),而HDL也并非仅有一种,而是有许多种,如ABEL(Advanced Boolean Expression Language)、AHDL(Altera Hardware Description Language)、Confluence、CUPL(Cornell University Programming Language)、HDCal、JHDL(Java Hardware Description Language)、Lava、Lola、MyHDL、PALASM、RHDL(Ruby Hardware Description Language)等,目前最普遍使用的是VHDL(Very-High-Speed Integrated Circuit Hardware Description Language)与Verilog。本领域技术人员也应该清楚,只需要将方法流程用上述几种硬件描述语言稍作逻辑编程并编程到集成电路中,就可以很容易得到实现该逻辑方法流程的硬件电路。

[0123] 控制器可以按任何适当的方式实现,例如,控制器可以采取例如微处理器或处理器以及存储可由该(微)处理器执行的计算机可读程序代码(例如软件或固件)的计算机可读介质、逻辑门、开关、专用集成电路(Application Specific Integrated Circuit,ASIC)、可编程逻辑控制器和嵌入微控制器的形式,控制器的例子包括但不限于以下微控制器:ARC 625D、Atmel AT91SAM、Microchip PIC18F26K20以及Silicone Labs C8051F320,存储器控制器还可以被实现为存储器的控制逻辑的一部分。本领域技术人员也知道,除了以纯计算机可读程序代码方式实现控制器以外,完全可以通过将方法步骤进行逻辑编程来使得控制器以逻辑门、开关、专用集成电路、可编程逻辑控制器和嵌入微控制器等的形式来实现相同功能。因此这种控制器可以被认为是一种硬件部件,而对其内包括的用于实现各种功能的装置也可以视为硬件部件内的结构。或者甚至,可以将用于实现各种功能的装置视为既可以是实现方法的软件模块又可以是硬件部件内的结构。

[0124] 上述实施例阐明的系统、装置、模块或单元,具体可以由计算机芯片或实体实现,或者由具有某种功能的产品来实现。一种典型的实现设备为计算机。具体的,计算机例如可以为个人计算机、膝上型计算机、蜂窝电话、相机电话、智能电话、个人数字符助理、媒体播放器、导航设备、电子邮件设备、游戏控制台、平板计算机、可穿戴设备或者这些设备中的任何设备的组合。

[0125] 为了描述的方便,描述以上装置时以功能分为各种单元分别描述。当然,在实施本说明书时可以把各单元的功能在同一个或多个软件和/或硬件中实现。

[0126] 本领域内的技术人员应明白,本发明的实施例可提供为方法、系统、或计算机程序产品。因此,本发明可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且,本发明可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0127] 本发明是参照根据本发明实施例的方法、设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流

程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器,使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

[0128] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中,使得存储在该计算机可读存储器中的指令产生包括指令装置的制品,该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

[0129] 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上,使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理,从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

[0130] 在一个典型的配置中,计算设备包括一个或多个处理器(CPU)、输入/输出接口、网络接口和内存。

[0131] 内存可能包括计算机可读介质中的非永久性存储器,随机存取存储器(RAM)和/或非易失性内存等形式,如只读存储器(ROM)或闪存(flash RAM)。内存是计算机可读介质的示例。

[0132] 计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括,但不限于相变内存(PRAM)、静态随机存取存储器(SRAM)、动态随机存取存储器(DRAM)、其他类型的随机存取存储器(RAM)、只读存储器(ROM)、电可擦除可编程只读存储器(EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器(CD-ROM)、数字多功能光盘(DVD)或其他光学存储、磁盒式磁带,磁带磁磁盘存储或其他磁性存储设备或任何其他非传输介质,可用于存储可以被计算设备访问的信息。按照本文中的界定,计算机可读介质不包括暂存电脑可读媒体(transitory media),如调制的数据信号和载波。

[0133] 还需要说明的是,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、商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、商品或者设备所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括所述要素的过程、方法、商品或者设备中还存在另外的相同要素。

[0134] 本领域技术人员应明白,本说明书的实施例可提供为方法、系统或计算机程序产品。因此,本说明书可采用完全硬件实施例、完全软件实施例或结合软件和硬件方面的实施例的形式。而且,本说明书可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0135] 本说明书可以在由计算机执行的计算机可执行指令的一般上下文中描述,例如程序模块。一般地,程序模块包括执行特定任务或实现特定抽象数据类型的例程、程序、对象、组件、数据结构等等。也可以在分布式计算环境中实践本说明书,在这些分布式计算环境中,由通过通信网络而被连接的远程处理设备来执行任务。在分布式计算环境中,程序模块

可以位于包括存储设备在内的本地和远程计算机存储介质中。

[0136] 以上所述仅为本说明书的实施例而已,并不用于限制本说明书。对于本领域技术人员来说,本说明书可以有各种更改和变化。凡在本说明书的精神和原理之内所作的任何修改、等同替换、改进等,均应包含在本说明书的权利要求范围之内。

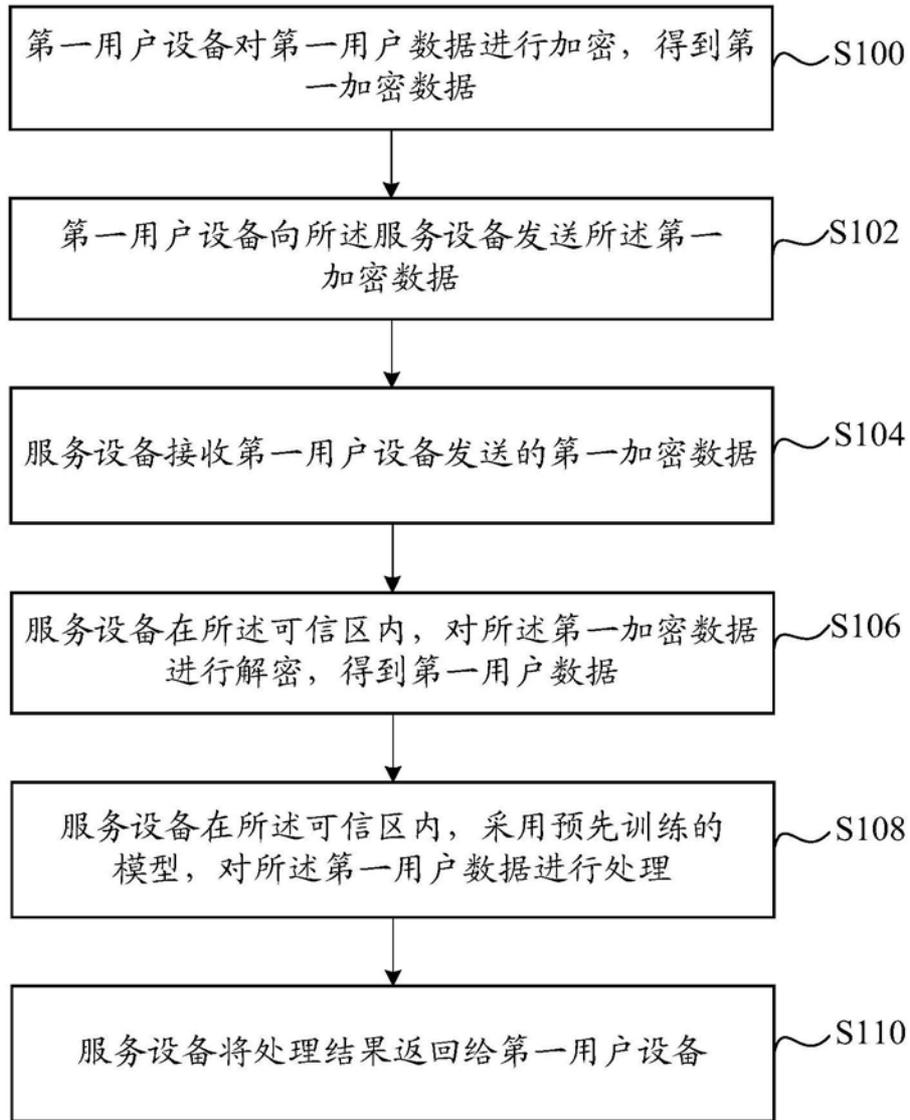


图1

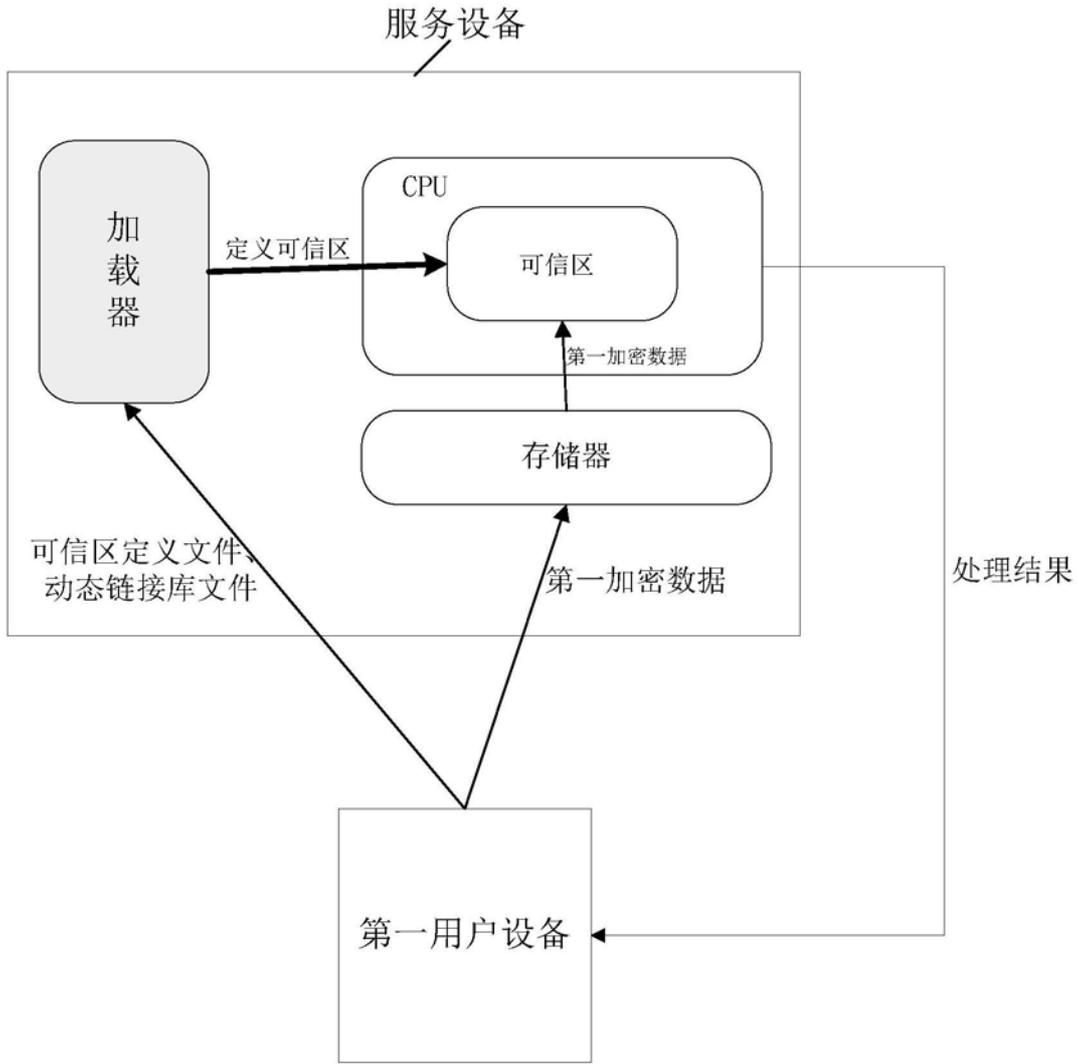


图2

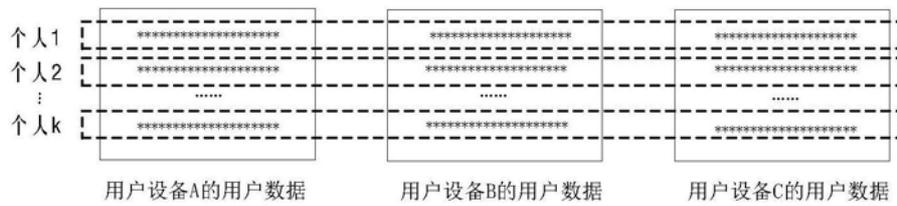


图3

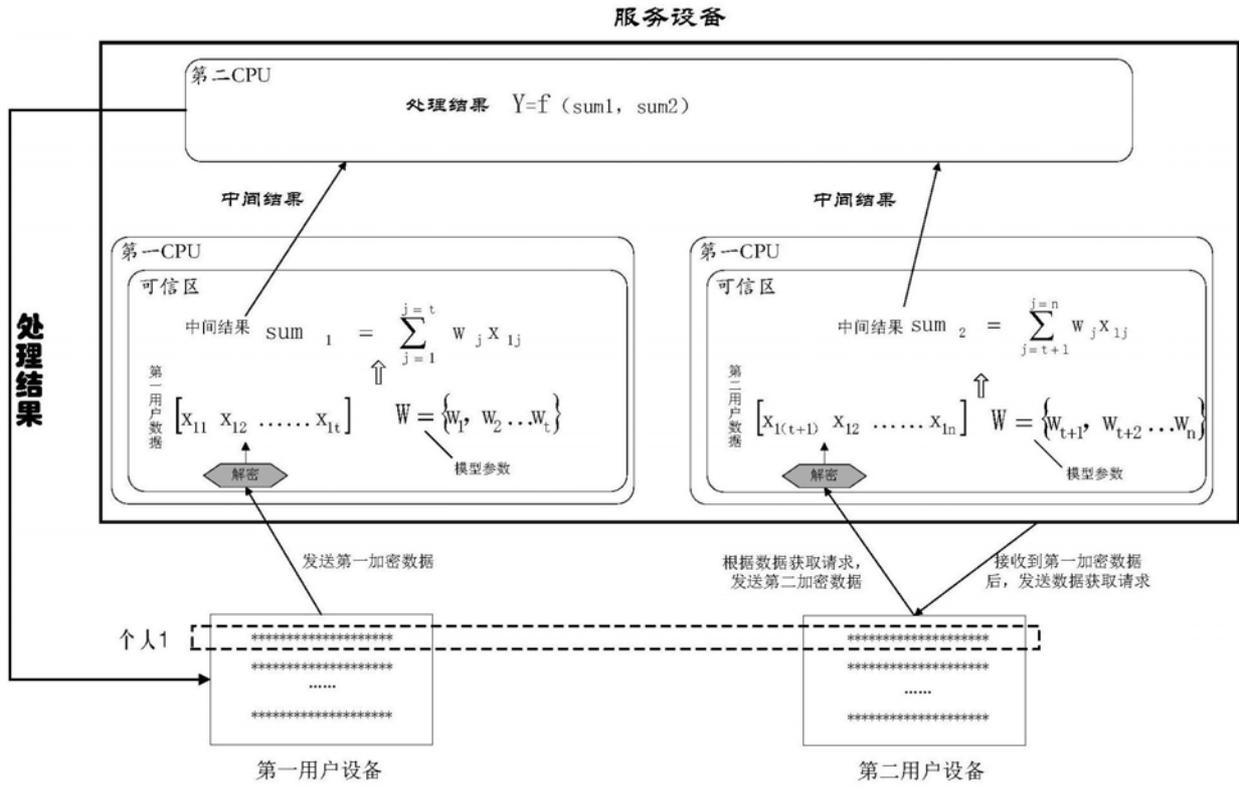


图4

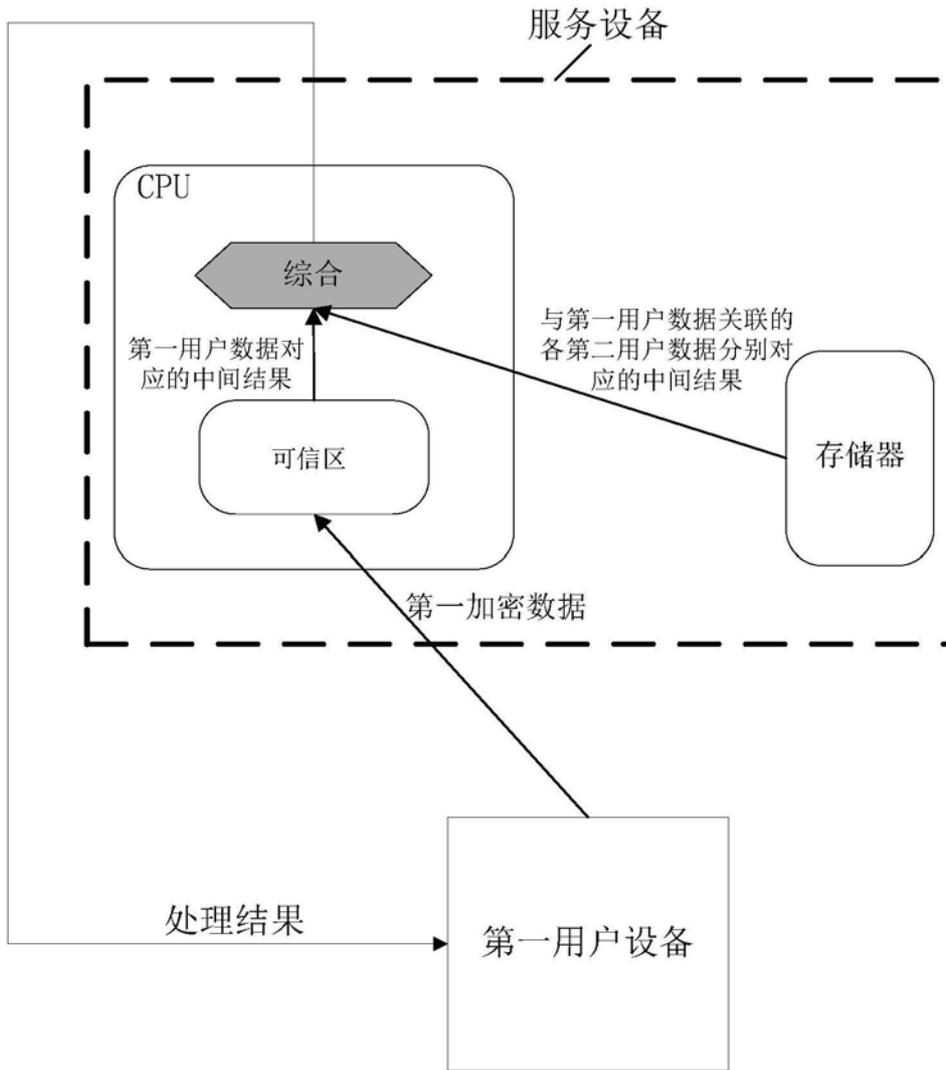


图5

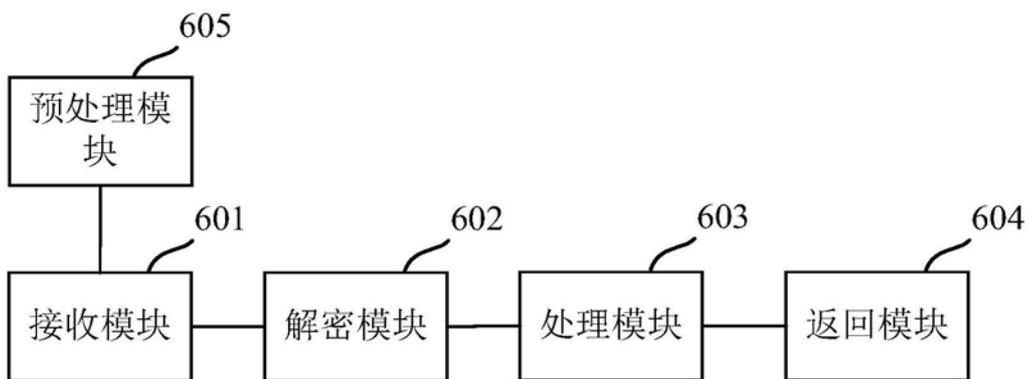


图6



图7

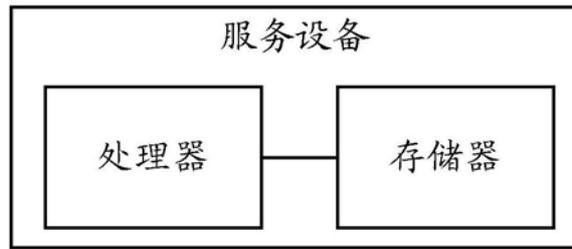


图8

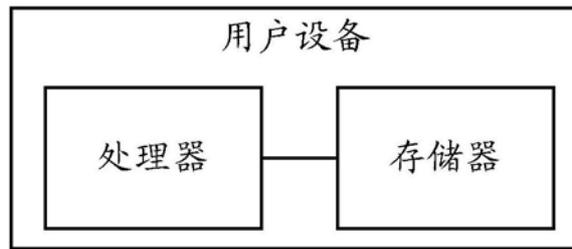


图9