(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 118382866 A
(43)申请公布日 2024.07.23

(21)申请号 202180104928.2
(22)申请日 2021.12.21
(85)PCT国际申请进入国家阶段日
2024.06.12
(86)PCT国际申请的申请数据
PCT/JP2021/047341 2021.12.21
(87)PCT国际申请的公布数据
WO2023/119421 JA 2023.06.29
(71)申请人 三菱电机株式会社
地址 日本东京都
(72)发明人 中井纲人
(74)专利代理机构 北京三友知识产权代理有限公司 11127
专利代理师 马建军 邓毅
(51)Int.Cl.
G06F 21/62(2006.01)

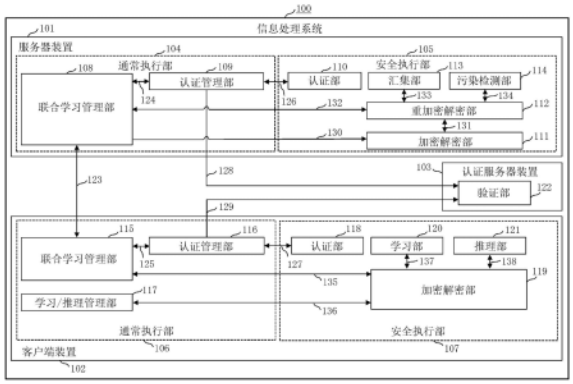
权利要求书4页 说明书19页 附图12页

(54)发明名称

信息处理系统、信息处理方法和信息处理程序

(57)摘要

服务器装置(101)和客户端装置(102)的各装置具有虚拟地分离的通常执行部和安全执行部。各装置的通常执行部彼此认证安全执行部的启动的正确性。在安全执行部的启动的正确性得到认证时,在各装置的安全执行部彼此之间建立安全的通信路径。服务器装置(101)的安全执行部对从客户端装置(102)经由安全的通信路径提供的模型信息进行解密并汇集。服务器装置(101)的安全执行部对通过汇集得到的模型信息进行加密后发送到服务器装置(101)的通常执行部。服务器装置(101)的通常执行部将通过汇集得到的模型信息在被加密的状态下储存于存储部。



CN 118382866 A

1.一种信息处理系统,其具有服务器装置和客户端装置,在所述服务器装置与所述客户端装置之间授受学习中使用的模型信息,其中,

作为虚拟地分离的执行环境,所述服务器装置和所述客户端装置的各装置具有通常的执行环境即通常执行部和安全的执行环境即安全执行部,

所述服务器装置和所述客户端装置的各装置的通常执行部彼此认证各装置的安全执行部的启动的正确性,在各装置的安全执行部的启动的正确性得到认证时,在各装置的安全执行部彼此之间建立收发被加密的数据的安全的通信路径,

所述服务器装置的安全执行部执行对从所述客户端装置经由所述安全的通信路径提供的所述模型信息进行解密并汇集的汇集处理,对通过汇集处理得到的模型信息进行加密后发送到所述服务器装置的通常执行部,

所述服务器装置的通常执行部将通过汇集处理得到的模型信息在被加密的状态下储存于存储部。

2.根据权利要求1所述的信息处理系统,其中,

所述模型信息包含从所述客户端装置向所述服务器装置提供的客户端模型和从所述服务器装置向所述客户端装置发布的全局模型,

所述服务器装置的安全执行部对从所述客户端装置经由所述安全的通信路径提供的客户端模型进行解密,对解密后的客户端模型进行重加密后发送到所述服务器装置的通常执行部,

所述服务器装置的通常执行部将被重加密的客户端模型储存于存储部。

3.根据权利要求2所述的信息处理系统,其中,

所述服务器装置的通常执行部将所述被重加密的客户端模型发送到所述服务器装置的安全执行部,

所述服务器装置的安全执行部对从所述服务器装置的通常执行部发送的客户端模型执行所述汇集处理,由此生成所述全局模型,对所述全局模型进行加密后发送到所述服务器装置的通常执行部,

所述服务器装置的通常执行部将被加密的全局模型储存于存储部。

4.根据权利要求2或3所述的信息处理系统,其中,

所述服务器装置的通常执行部对所述被重加密的客户端模型进行分割后发送到所述服务器装置的安全执行部,

所述服务器装置的安全执行部按照被分割后的每个客户端模型执行汇集处理。

5.根据权利要求2~4中的任意一项所述的信息处理系统,其中,

所述服务器装置的安全执行部执行检测所述解密后的客户端模型是否被污染的污染检测处理,不汇集被污染的客户端模型。

6.根据权利要求2~4中的任意一项所述的信息处理系统,其中,

所述客户端装置的安全执行部执行检测向所述服务器装置提供的客户端模型是否被污染的污染检测处理,被污染的客户端模型不提供给所述服务器装置。

7.一种信息处理系统,其具有服务器装置和客户端装置,在所述服务器装置与所述客户端装置之间授受学习中使用的模型信息,其中,

作为虚拟地分离的执行环境,所述客户端装置具有通常的执行环境即通常执行部和安

全的执行环境即安全执行部，

所述服务器装置仅具有通常的执行环境即通常执行部，

所述客户端装置的安全执行部对向所述服务器装置提供的模型信息执行同态加密，

所述服务器装置的通常执行部对被同态加密的模型信息执行在同态加密的状态下进行汇集的汇集处理，将通过汇集处理得到的模型信息在被同态加密的状态下储存于存储部。

8.根据权利要求7所述的信息处理系统，其中，

所述服务器装置的通常执行部对被同态加密的模型信息执行在同态加密的状态下检测污染的污染检测处理。

9.根据权利要求7所述的信息处理系统，其中，

所述客户端装置的安全执行部执行检测向所述服务器装置提供的模型信息是否被污染的污染检测处理，被污染的模型信息不提供给所述服务器装置。

10.一种信息处理系统，其具有服务器装置和客户端装置，在所述服务器装置与所述客户端装置之间授受学习中使用的模型信息，其中，

作为虚拟地分离的执行环境，所述服务器装置和所述客户端装置的各装置具有通常的执行环境即通常执行部和安全的执行环境即安全执行部，

所述服务器装置和所述客户端装置的各装置的通常执行部彼此认证各装置的安全执行部的启动的正确性，在各装置的安全执行部的启动的正确性得到认证时，在各装置的安全执行部彼此之间建立收发被加密的数据的安全的通信路径，

所述服务器装置的安全执行部对从所述客户端装置经由所述安全的通信路径提供的所述模型信息执行同态加密，将被同态加密的模型信息在被同态加密的状态下储存于存储部，

所述服务器装置的通常执行部对被同态加密的模型信息执行在同态加密的状态下进行汇集的汇集处理，将通过汇集处理得到的模型信息在被同态加密的状态下储存于存储部。

11.根据权利要求10所述的信息处理系统，其中，

所述服务器装置的通常执行部对被同态加密的模型信息执行在同态加密的状态下检测污染的污染检测处理。

12.根据权利要求10所述的信息处理系统，其中，

所述客户端装置的安全执行部执行检测向所述服务器装置提供的模型信息是否被污染的污染检测处理，被污染的模型信息不提供给所述服务器装置。

13.一种信息处理系统中使用的信息处理方法，该信息处理系统具有服务器装置和客户端装置，在所述服务器装置与所述客户端装置之间授受学习中使用的模型信息，其中，

作为虚拟地分离的执行环境，所述服务器装置和所述客户端装置的各装置具有通常的执行环境和安全的执行环境，

在所述服务器装置和所述客户端装置的各装置的通常的执行环境中，彼此认证各装置的安全的执行环境的启动的正确性，在各装置的安全的执行环境的启动的正确性得到认证时，在各装置的安全的执行环境彼此之间建立收发被加密的数据的安全的通信路径，

在所述服务器装置的安全的执行环境中，执行对从所述客户端装置经由所述安全的通

信路径提供的所述模型信息进行解密并汇集的汇集处理,对通过汇集处理得到的模型信息进行加密后发送到所述服务器装置的通常的执行环境,

在所述服务器装置的通常的执行环境中,将通过汇集处理得到的模型信息在被加密的状态下储存于存储器。

14.一种信息处理系统中使用的信息处理方法,该信息处理系统具有服务器装置和客户端装置,在所述服务器装置与所述客户端装置之间授受学习中使用的模型信息,其中,

作为虚拟地分离的执行环境,所述客户端装置具有通常的执行环境和安全的执行环境,

所述服务器装置仅具有通常的执行环境,

在所述客户端装置的安全的执行环境中,对向所述服务器装置提供的模型信息执行同态加密,

在所述服务器装置的通常的执行环境中,对被同态加密的模型信息执行在同态加密的状态下进行汇集的汇集处理,将通过汇集处理得到的模型信息在被同态加密的状态下储存于存储器。

15.一种信息处理系统中使用的信息处理方法,该信息处理系统具有服务器装置和客户端装置,在所述服务器装置与所述客户端装置之间授受学习中使用的模型信息,其中,

作为虚拟地分离的执行环境,所述服务器装置和所述客户端装置的各装置具有通常的执行环境和安全的执行环境,

在所述服务器装置和所述客户端装置的各装置的通常的执行环境中,彼此认证各装置的安全的执行环境的启动的正确性,在各装置的安全的执行环境的启动的正确性得到认证时,在各装置的安全的执行环境彼此之间建立收发被加密的数据的安全的通信路径,

在所述服务器装置的安全的执行环境中,对从所述客户端装置经由所述安全的通信路径提供的所述模型信息执行同态加密,将被同态加密的模型信息在被同态加密的状态下储存于存储器,

在所述服务器装置的通常的执行环境中,对被同态加密的模型信息执行在同态加密的状态下进行汇集的汇集处理,将通过汇集处理得到的模型信息在被同态加密的状态下储存于存储器。

16.一种信息处理系统中使用的信息处理程序,该信息处理系统具有服务器装置和客户端装置,在所述服务器装置与所述客户端装置之间授受学习中使用的模型信息,其中,

作为虚拟地分离的执行环境,所述服务器装置和所述客户端装置的各装置具有通常的执行环境和安全的执行环境,

所述信息处理程序使计算机执行以下处理:

在所述服务器装置和所述客户端装置的各装置的通常的执行环境中,彼此认证各装置的安全的执行环境的启动的正确性,在各装置的安全的执行环境的启动的正确性得到认证时,在各装置的安全的执行环境彼此之间建立收发被加密的数据的安全的通信路径;

在所述服务器装置的安全的执行环境中,执行对从所述客户端装置经由所述安全的通信路径提供的所述模型信息进行解密并汇集的汇集处理,对通过汇集处理得到的模型信息进行加密后发送到所述服务器装置的通常的执行环境;以及

在所述服务器装置的通常的执行环境中,将通过汇集处理得到的模型信息在被加密的

状态下储存于存储器。

17.一种信息处理系统中使用的信息处理程序,该信息处理系统具有服务器装置和客户端装置,在所述服务器装置与所述客户端装置之间授受学习中使用的模型信息,其中,

作为虚拟地分离的执行环境,所述客户端装置具有通常的执行环境和安全的执行环境,

所述服务器装置仅具有通常的执行环境,

所述信息处理程序使计算机执行以下处理:

在所述客户端装置的安全的执行环境中,对向所述服务器装置提供的模型信息执行同态加密;以及

在所述服务器装置的通常的执行环境中,对被同态加密的模型信息执行在同态加密的状态下进行汇集的汇集处理,将通过汇集处理得到的模型信息在被同态加密的状态下储存于存储器。

18.一种信息处理系统中使用的信息处理程序,该信息处理系统具有服务器装置和客户端装置,在所述服务器装置与所述客户端装置之间授受学习中使用的模型信息,其中,

作为虚拟地分离的执行环境,所述服务器装置和所述客户端装置的各装置具有通常的执行环境和安全的执行环境,

所述信息处理程序使计算机执行以下处理:

在所述服务器装置和所述客户端装置的各装置的通常的执行环境中,彼此认证各装置的安全的执行环境的启动的正确性,在各装置的安全的执行环境的启动的正确性得到认证时,在各装置的安全的执行环境彼此之间建立收发被加密的数据的安全的通信路径;

在所述服务器装置的安全的执行环境中,对从所述客户端装置经由所述安全的通信路径提供的所述模型信息执行同态加密,将被同态加密的模型信息在被同态加密的状态下储存于存储器;以及

在所述服务器装置的通常的执行环境中,对被同态加密的模型信息执行在同态加密的状态下进行汇集的汇集处理,将通过汇集处理得到的模型信息在被同态加密的状态下储存于存储器。

信息处理系统、信息处理方法和信息处理程序

技术领域

[0001] 本发明涉及信息处理系统、信息处理方法和信息处理程序。特别涉及作为以联合学习为代表的分布式机器学习系统的信息处理系统、信息处理方法和信息处理程序。

背景技术

[0002] 现有的以联合学习为代表的分布式机器学习系统不是将学习数据从客户端收集到服务器,而是将在客户端进行了学习的模型即客户端模型收集到服务器。由此,进行考虑到客户端的学习数据中包含的隐私信息的模型学习。但是,开始被指出从客户端模型泄露客户端的学习数据中包含的隐私信息。

[0003] 在非专利文献1中,使用服务器上的安全的执行环境即TEE,在TEE内对客户端模型进行处理,由此进行考虑到隐私信息的模型学习。TEE是Trusted Execution Environment(可信执行环境)的简称。

[0004] 现有技术文献

[0005] 非专利文献

[0006] 非专利文献1:L.Zhao et al.,“SEAR:Secure and Efficient Aggregation for Byzantine-Robust Federated Learning,”IEEE Transactions on Dependable and Secure Computing,2021

发明内容

[0007] 发明要解决的课题

[0008] 在现有的以联合学习为代表的分布式机器学习系统中,存在以下3个主要的安全/隐私的问题。

[0009] (1)来自从设备或边缘发送的客户端模型的隐私信息泄露的问题

[0010] (2)由于来自存在恶意的设备或边缘的伪信息而引起的学习污染和干扰的问题

[0011] (3)全局模型的窃取或复制的问题

[0012] 但是,在现有的考虑到安全/隐私的技术中,没有解决上述3个问题全部。例如,在非专利文献1中,针对上述的问题(1)和(2),提出使用TEE这样的安全的执行环境的解决方案。但是,没有公开针对问题(3)的解决方案。进而,在非专利文献1中,存在由于使用安全的执行环境而引起的针对系统的负荷变大这样的课题。

[0013] 在本发明中,其目的在于,提供抑制安全对策对系统的负荷并且实现考虑到安全/隐私的联合学习的信息处理系统。

[0014] 用于解决课题的手段

[0015] 本发明的信息处理系统具有服务器装置和客户端装置,在所述服务器装置与所述客户端装置之间授受学习中使用的模型信息,其中,作为虚拟地分离的执行环境,所述服务器装置和所述客户端装置的各装置具有通常的执行环境即通常执行部和安全的执行环境即安全执行部,所述服务器装置和所述客户端装置的各装置的通常执行部彼此认证各装置

的安全执行部的启动的正确性,在各装置的安全执行部的启动的正确性得到认证时,在各装置的安全执行部彼此之间建立收发被加密的数据的安全的通信路径,所述服务器装置的安全执行部执行对从所述客户端装置经由所述安全的通信路径提供的模型信息进行解密并汇集的汇集处理,对通过汇集处理得到的模型信息进行加密后发送到所述服务器装置的通常执行部,所述服务器装置的通常执行部将通过所述汇集处理得到的模型信息在被加密的状态下储存于存储部。

[0016] 发明效果

[0017] 在本发明的信息处理系统中,服务器装置的安全执行部对从客户端装置经由安全的通信路径提供的模型信息进行解密并汇集。然后,服务器装置的安全执行部对通过汇集得到的模型信息进行加密后发送到服务器装置的通常执行部。服务器装置的通常执行部将通过汇集得到的模型信息在被加密的状态下储存于存储部。由此,根据本发明的信息处理系统,能够提供抑制安全对策对系统的负荷并且实现考虑到安全/隐私的联合学习的信息处理系统。

附图说明

[0018] 图1是示出实施方式1的信息处理系统的结构例的图。

[0019] 图2是示出实施方式1的服务器装置的硬件结构例的图。

[0020] 图3是示出实施方式1的信息处理系统中的客户端模型收集的动作的序列图。

[0021] 图4是示出实施方式1的信息处理系统中的全局模型发布的动作的序列图。

[0022] 图5是示出实施方式1的变形例的信息处理系统的硬件结构例的图。

[0023] 图6是示出实施方式2的信息处理系统的结构例的图。

[0024] 图7是示出实施方式2的信息处理系统中的客户端模型收集的动作的序列图。

[0025] 图8是示出实施方式2的信息处理系统中的全局模型发布的动作的序列图。

[0026] 图9是示出实施方式3的信息处理系统的结构例的图。

[0027] 图10是示出实施方式3的信息处理系统中的客户端模型收集的动作的序列图。

[0028] 图11是示出实施方式3的信息处理系统中的全局模型发布的动作的序列图。

[0029] 图12是示出实施方式4的信息处理系统的结构例的图。

具体实施方式

[0030] 下面,使用附图对本实施方式进行说明。在各图中,对相同或相当的部分标注相同标号。在实施方式的说明中,关于相同或相当的部分,适当省略或简化说明。

[0031] 实施方式1

[0032] ***结构的说明***

[0033] 图1是示出本实施方式的信息处理系统100的结构例的图。

[0034] 信息处理系统100具有服务器装置101、客户端装置102和认证服务器装置103。客户端装置102存在多个。服务器装置也称作服务器部。客户端装置也称作客户端部。认证服务器装置也称作认证服务器部。

[0035] 信息处理系统100在服务器装置101与客户端装置102之间授受在学习中使用的模型信息。

[0036] 模型信息包含客户端模型和全局模型。客户端模型是从客户端装置102向服务器装置101提供的学习模型。全局模型是从服务器装置101向客户端装置102发布的学习模型。全局模型是通过汇集从客户端装置102收集的客户端模型而生成的。

[0037] 服务器装置101、客户端装置102和认证服务器装置103分别是计算机，经由网络进行信息的交换。

[0038] 另外，服务器装置101、客户端装置102和认证服务器装置103也可以分别搭载于独立的计算机。或者，服务器装置101、客户端装置102和认证服务器装置103也可以搭载于1个计算机，虚拟地构成3个计算机。或者，服务器装置101、客户端装置102和认证服务器装置103中的服务器装置101和认证服务器装置103这样的一部分也可以搭载于1个计算机，虚拟地构成多个计算机。

[0039] 在以下的说明中，有时将服务器装置101、客户端装置102和认证服务器装置103分别称作信息处理系统100的各装置。

[0040] 信息处理系统100的各装置是计算机。信息处理系统100的各装置具有处理器，并且具有存储器、辅助存储装置、输入接口、输出接口和通信装置这样的其他硬件。处理器经由信号线与其他硬件连接，对这些其他硬件进行控制。

[0041] 作为虚拟地分离的执行环境，服务器装置101和客户端装置102的各装置具有通常的执行环境即通常执行部和安全的执行环境即安全执行部。虚拟地分离的执行环境在后面说明。

[0042] 作为功能要素，服务器装置101具有通常执行部104和安全执行部105。通常执行部104具有联合学习管理部108和认证管理部109。安全执行部105具有认证部110、加密解密部111、重加密解密部112、汇集部113和污染检测部114。

[0043] 另外，虽然没有图示，但是，通常执行部104和安全执行部105分别具有存储部。在存储部中存储有在信息处理中使用的客户端模型、全局模型、秘钥和认证信息这样的信息。

[0044] 在以下的说明中，在记载为“储存于通常执行部”或“存储于通常执行部”的情况下，意味着“储存于分配给通常执行部的存储部”或“存储于分配给通常执行部的存储部”。此外，在记载为“储存于安全执行部”或“存储于安全执行部”的情况下，意味着“储存于分配给安全执行部的存储部”或“存储于分配给安全执行部的存储部”。在以下的客户端装置102和认证服务器装置103中也同样。

[0045] 作为功能要素，客户端装置102具有通常执行部106和安全执行部107。通常执行部106具有联合学习管理部115、认证管理部116和学习/推理管理部117。安全执行部107具有认证部118、加密解密部119、学习部120和推理部121。

[0046] 另外，虽然没有图示，但是，通常执行部106和安全执行部107分别具有存储部。在存储部中存储有在信息处理中使用的客户端模型、全局模型、秘钥和认证信息这样的信息。

[0047] 作为功能要素，认证服务器装置103具有验证部122。

[0048] 另外，虽然没有图示，但是，认证服务器装置103具有存储部。在存储部中存储有由验证部122验证的认证信息这样的信息。

[0049] 图2是示出本实施方式的服务器装置101的硬件结构例的图。

[0050] 以图2的服务器装置101为例对信息处理系统100的各装置的硬件结构例进行说明。客户端装置102和认证服务器装置103的硬件结构例与服务器装置101相同，因此省略图

示。

[0051] 服务器装置101是计算机。服务器装置101具有处理器910,并且具有存储器921、辅助存储装置922、输入接口930、输出接口940和通信装置950这样的其他硬件。处理器910经由信号线与其他硬件连接,对这些其他硬件进行控制。

[0052] 在服务器装置101中,通常执行部104和安全执行部105的功能通过软件来实现。存储部设置于存储器921。另外,存储部也可以设置于辅助存储装置922,还可以分散设置于存储器921和辅助存储装置922。

[0053] 处理器910是在服务器装置101中执行信息处理程序的装置。信息处理程序是实现信息处理系统100的各装置的功能的程序。

[0054] 处理器910是进行运算处理的IC。处理器910的具体例是CPU、DSP或GPU。IC是Integrated Circuit(集成电路)的简称。CPU是Central Processing Unit(中央处理单元)的简称。DSP是Digital Signal Processor(数字信号处理器)的简称。GPU是Graphics Processing Unit(图形处理单元)的简称。

[0055] 存储器921是临时存储数据的存储装置。存储器921的具体例是SRAM或DRAM。SRAM是Static Random Access Memory(静态随机存取存储器)的简称。DRAM是Dynamic Random Access Memory(动态随机存取存储器)的简称。

[0056] 辅助存储装置922是保管数据的存储装置。辅助存储装置922的具体例是HDD。此外,辅助存储装置922也可以是SD(注册商标)存储卡、CF、NAND闪存、软盘、光盘、高密度盘、蓝光(注册商标)盘、DVD这样的移动存储介质。另外,HDD是Hard Disk Drive(硬盘驱动器)的简称。SD(注册商标)是Secure Digital(安全数字)的简称。CF是CompactFlash(致密闪存)(注册商标)的简称。DVD是Digital Versatile Disk(数字多功能盘)的简称。

[0057] 输入接口930是与鼠标、键盘或触摸面板这样的输入装置连接的端口。具体而言,输入接口930是USB端子。另外,输入接口930也可以是与LAN连接的端口。USB是Universal Serial Bus(通用串行总线)的简称。LAN是Local Area Network(局域网)的简称。

[0058] 输出接口940是连接显示器这样的输出设备的缆线的端口。具体而言,输出接口940是USB端子或HDMI(注册商标)端子。具体而言,显示器是LCD。输出接口940也称作显示器接口。HDMI(注册商标)是High Definition Multimedia Interface(高清晰度多媒体接口)的简称。LCD是Liquid Crystal Display(液晶显示器)的简称。

[0059] 通信装置950具有接收机和发送机。通信装置950与LAN、互联网或电话线路这样的通信网连接。具体而言,通信装置950是通信芯片或NIC。NIC是Network Interface Card(网络接口卡)的简称。

[0060] 信息处理程序在服务器装置101中执行。信息处理程序被读入到处理器910,由处理器910来执行。在存储器921中,不仅存储有信息处理程序,还存储有OS(Operating System:操作系统)。处理器910一边执行OS,一边执行信息处理程序。信息处理程序和OS也可以存储于辅助存储装置922。辅助存储装置922中存储的信息处理程序和OS被载入到存储器921,由处理器910来执行。另外,信息处理程序的一部分或全部也可以嵌入OS中。

[0061] 服务器装置101也可以具有代替处理器910的多个处理器。这些多个处理器分担执行信息处理程序。与处理器910同样,各个处理器是执行信息处理程序的装置。

[0062] 由信息处理程序利用、处理或输出的数据、信息、信号值和变量值存储于存储器

921、辅助存储装置922或处理器910内的寄存器或高速缓冲存储器。

[0063] 也可以将通常执行部104和安全执行部105的各部的“部”改写成“电路”、“工序”、“步骤”、“处理”或“线路”。信息处理程序使计算机执行通常执行处理和安全执行处理。也可以将通常执行处理和安全执行处理的“处理”改写成“程序”、“程序产品”、“存储有程序的计算机能读取的存储介质”或“记录有程序的计算机能读取的记录介质”。此外,信息处理方法是信息处理系统100的各装置执行信息处理程序而实现的方法。

[0064] 信息处理程序也可以储存于计算机能读取的记录介质来提供。此外,信息处理程序也可以作为程序产品来提供。

[0065] ***功能的说明***

[0066] 接着,使用图1对信息处理系统100的各装置的功能进行说明。

[0067] 图1所示的信息处理系统100在由服务器装置101和客户端装置102构成的以联合学习为代表的分布式机器学习系统中的信息处理系统中追加了认证服务器装置103。

[0068] 作为虚拟地分离的执行环境,服务器装置101和客户端装置102的各装置具有通常的执行环境即通常执行部和安全的执行环境即安全执行部。

[0069] 服务器装置101能够虚拟地分离成通常执行部104和安全执行部105。

[0070] 在服务器装置101中,在通常执行部104中具有联合学习管理部108和认证管理部109。

[0071] 联合学习管理部108对以联合学习为代表的分布式机器学习的执行进行管理。

[0072] 认证管理部109验证安全执行部105的正确性。

[0073] 此外,在服务器装置101中,在安全执行部105中具有认证部110、加密解密部111、重加密解密部112、汇集部113和污染检测部114。

[0074] 认证部110提供用于验证安全执行部105的正确性的认证信息。

[0075] 加密解密部111对与客户端装置102交换的模型信息进行加密或解密处理。与客户端装置102交换的模型信息是客户端模型和全局模型。

[0076] 重加密解密部112对与通常执行部104交换的信息进行重加密或解密处理。

[0077] 汇集部113汇集客户端模型。

[0078] 污染检测部114检测客户端模型的污染。

[0079] 客户端装置102能够虚拟地分离成通常执行部106和安全执行部107。

[0080] 在客户端装置102中,在通常执行部106中具有联合学习管理部115、认证管理部116和学习/推理管理部117。

[0081] 联合学习管理部115对以联合学习为代表的分布式机器学习的执行进行管理。

[0082] 认证管理部116验证安全执行部107的正确性。

[0083] 学习/推理管理部117对模型信息的学习和推理的执行进行管理。

[0084] 此外,在客户端装置102中,在安全执行部107中具有认证部118、加密解密部119、学习部120和推理部121。

[0085] 认证部118提供用于验证安全执行部107的正确性的认证信息。

[0086] 加密解密部119对与服务器装置101交换的模型信息进行加密或解密处理。与服务器装置101交换的模型信息是客户端模型或全局模型。

[0087] 学习部120执行模型信息的学习。

[0088] 推理部121使用模型信息来执行推理。

[0089] 认证服务器装置103具有验证部122。

[0090] 验证部122验证安全执行部105和安全执行部107各自的认证信息。

[0091] 下面,例如,有时对服务器装置101的安全执行部105和客户端装置102的安全执行部107进行说明。此时,有时如安全执行部105和107、安全执行部105或107、或者安全执行部105、107那样省略结构要素名。

[0092] 图1的信息处理系统100通过设为上述这种结构,保护客户端模型和全局模型,检测安全执行部105和107各自的正确性以及客户端模型的污染。由此,实现考虑到安全/隐私的联合学习。

[0093] ***功能的详细说明***

[0094] 接着,使用图1更加详细地说明信息处理系统100的各装置的功能。

[0095] 以联合学习为代表的分布式机器学习算法通过服务器装置101和客户端装置102各自的联合学习管理部108、115彼此的交换123来执行。设想客户端装置102存在多个。

[0096] 通常执行部104、106以及安全执行部105、107的虚拟分离例如通过Arm Trustzone或Intel(注册商标)SGX这样的TEE技术来实现。

[0097] 联合学习管理部108、115进行联合学习用的客户端模型的收集或全局模型的发布。此外,联合学习管理部108、115通过认证管理部109、116验证彼此的安全执行部105、107的正确性(处理124、125)。

[0098] 另外,在图1中,对结构要素之间的箭头赋予编号。该箭头表示结构要素之间的交换。在以下的说明中,将该箭头所示的交换称作“处理”。在以下的图6、图9和图12中也同样。

[0099] 认证管理部109、116从位于安全执行部105、107的认证部110、118取得用于验证安全执行部105、107的正确性的认证信息(处理126、127)。

[0100] 认证部110、118输出认证信息(处理126、127)。认证信息例如是已启动的安全执行部的哈希值和签名。安全执行部105、107的认证例如通过Remote Attestation技术来实现。

[0101] 对认证服务器装置103的功能要素进行说明。

[0102] 验证部122取得分别来自认证管理部109、116的认证信息,验证安全执行部105、107分别是否正确地启动(处理128、129)。

[0103] 对服务器装置101的功能要素进行说明。

[0104] 加密解密部111按照每个客户端,对通过联合学习管理部108从客户端装置102收集到的客户端模型进行解密处理(处理130)。或者,加密解密部111按照每个客户端,对通过联合学习管理部108向客户端装置102发布的全局模型进行加密处理(处理130)。

[0105] 重加密解密部112利用临时的公共秘钥对收集到的客户端模型进行重加密处理(处理131、132),将其储存于通常执行部104的存储部。或者,重加密解密部112从通常执行部104取得被重加密的客户端模型,进行解密处理(处理132)。

[0106] 汇集部113取得收集到的已解密的客户端模型(处理133),进行汇集。汇集例如是指计算客户端模型的平均值。

[0107] 污染检测部114取得收集到的已解密的客户端模型(处理134),进行客户端模型的污染检测。污染检测例如是指,计算客户端模型之间的模型间距离,在距离大的情况下,检测为该客户端模型被污染。

[0108] 对客户端装置102的功能要素进行说明。

[0109] 加密解密部119进行通过联合学习管理部115向服务器装置101提供的客户端模型的加密处理(处理135)。或者,加密解密部119对通过联合学习管理部115从服务器装置101发布的全局模型进行解密处理(处理136)。

[0110] 学习/推理管理部117使用从服务器装置101发布的全局模型,对学习或推理处理的执行进行管理(处理136)。

[0111] 学习部120通过学习/推理管理部117,使用由加密解密部119解密后的全局模型(处理137)来执行学习。

[0112] 推理部121通过学习/推理管理部117,使用由加密解密部119解密后的全局模型(处理138)来执行推理。

[0113] 用于执行机器学习运算的学习/推理管理部117、学习部120和推理部121不限于深度学习。例如,学习/推理管理部117、学习部120和推理部121也可以是使用回归法、决策树学习、贝叶斯法或聚类这样的方法的运算。

[0114] ***动作的说明***

[0115] 接着,对本实施方式的信息处理系统100的动作进行说明。信息处理系统100的动作步骤相当于信息处理方法。此外,实现信息处理系统100的动作的程序相当于信息处理程序。

[0116] 图3是示出本实施方式的信息处理系统100中的客户端模型收集的动作的序列图。

[0117] 图4是示出本实施方式的信息处理系统100中的全局模型发布的动作的序列图。

[0118] 该序列图分成通常执行部104、106以及安全执行部105、107来示出信息处理系统100中的服务器装置101和客户端装置102的交换。

[0119] <客户端模型收集>

[0120] 使用图3对信息处理系统100中的客户端模型收集处理的动作进行说明。

[0121] 首先,服务器装置101和客户端装置102的各装置的通常执行部104、106彼此认证各装置的安全执行部的启动的正确性。在各装置的安全执行部的启动的正确性得到认证时,在各装置的安全执行部彼此之间建立收发被加密的数据的安全的通信路径。即,在各装置的安全的执行环境彼此之间建立安全的通信路径。

[0122] 具体而言,如下所述。

[0123] 在步骤S101中,服务器装置101的通常执行部104向客户端装置102的通常执行部106发送客户端模型的提供委托。

[0124] 在步骤S102中,客户端装置102的通常执行部106向服务器装置101的通常执行部104发送安全执行部的认证委托,以验证服务器装置101的安全执行部105的正确性。

[0125] 在步骤S103中,服务器装置101的通常执行部104向服务器装置101的安全执行部105发送认证信息的提供委托。

[0126] 在步骤S104中,服务器装置101的安全执行部105向服务器装置101的通常执行部104发送认证信息和公开秘钥PKs。

[0127] 在步骤S105中,服务器装置101的通常执行部104向客户端装置102的通常执行部106转发认证信息和公开秘钥PKs。客户端装置102的通常执行部106向认证服务器装置103的验证部122发送认证信息的验证委托。认证服务器装置103的验证部122向客户端装置102

的通常执行部106发送验证结果。客户端装置102的通常执行部106在能够验证服务器装置101的安全执行部105的正确性的情况下,向客户端装置102的安全执行部107发送公开秘钥PKs。

[0128] 在步骤S106中,客户端装置102的安全执行部107使用公开秘钥PKs与服务器装置101的安全执行部105进行秘钥交换,建立收发数据被加密的安全的通信路径。

[0129] 在步骤S107中,客户端装置102的安全执行部107在安全的通信路径上向服务器装置101的安全执行部105发送客户端模型M。

[0130] 这里,服务器装置101如以下那样进行动作,以抑制安全执行部105的消耗存储器。

[0131] 服务器装置101的安全执行部105对从客户端装置102经由安全的通信路径提供的客户端模型进行解密。然后,服务器装置101的安全执行部105对解密后的客户端模型进行重加密后发送到服务器装置101的通常执行部104。

[0132] 服务器装置101的通常执行部104将被重加密的客户端模型储存于存储部。

[0133] 具体而言,如下所述。

[0134] 在步骤S108中,服务器装置101的安全执行部105利用运算用的临时秘钥MKs对客户端模型M进行重加密。具体而言,服务器装置101的安全执行部105对在步骤S107中从客户端装置102接收到的客户端模型M进行解密,利用运算用的临时秘钥MKs进行重加密。然后,服务器装置101的安全执行部105向服务器装置101的通常执行部104发送利用临时秘钥MKs重加密后的客户端模型EncMKs(M)。服务器装置101的通常执行部104将利用临时秘钥MKs重加密后的客户端模型EncMKs(M)储存于存储部。

[0135] 通过该步骤S108的处理,服务器装置101能够抑制安全执行部105的消耗存储器。

[0136] 在信息处理系统100中,在各客户端装置102中执行步骤S101～步骤S108的处理,从各客户端装置102收集客户端模型。在全部客户端模型的收集完成后,进入接下来的步骤。

[0137] <客户端模型的汇集>

[0138] 接着,在服务器装置101中,进行客户端模型的汇集,生成全局模型。

[0139] 服务器装置101的安全执行部105执行对从客户端装置102经由安全的通信路径提供的模型信息进行解密并汇集的汇集处理。

[0140] 服务器装置101的安全执行部105对通过汇集处理得到的模型信息进行加密后发送到服务器装置101的通常执行部104。这里,模型信息是客户端模型。

[0141] 服务器装置101的通常执行部104将通过汇集处理得到的模型信息作为全局模型,在被加密的状态下储存于存储部。

[0142] 具体而言,如下所述。

[0143] 接着,在步骤S109中,服务器装置101的通常执行部104将被重加密的客户端模型EncMKs(M)发送到服务器装置101的安全执行部105。具体而言,服务器装置101的通常执行部104对被重加密的全部客户端模型EncMKs(M)进行分割后发送到服务器装置101的安全执行部105。服务器装置101的安全执行部105将被重加密的全部客户端模型EncMKs(M)分割成若干个并进行发送。

[0144] 服务器装置101的安全执行部105对被分割后的客户端模型EncMKs(M)进行解密。服务器装置101的安全执行部105存储被解密的客户端模型DecMKs(M)。安全执行部105中存

储的客户端模型DecMKs(M)是全部客户端模型的一部分。

[0145] 通过该步骤S109的处理,服务器装置101能够抑制安全执行部105的消耗存储器。

[0146] 服务器装置101的安全执行部105对从服务器装置101的通常执行部104发送的客户端模型执行汇集处理,由此生成全局模型。此时,服务器装置101的安全执行部105对客户端模型进行污染检测处理,不汇集被检测到污染的客户端模型。

[0147] 具体而言,如下所述。

[0148] 在步骤S110中,服务器装置101的安全执行部105使用被解密的客户端模型DecMKs(M)执行污染检测和汇集。

[0149] 服务器装置101的安全执行部105按照被分割后的每个客户端模型DecMKs(M)执行汇集处理。

[0150] 此外,服务器装置101的安全执行部105执行检测被解密的客户端模型DecMKs(M)是否被污染的污染检测处理。然后,服务器装置101的安全执行部105不汇集被检测到污染的客户端模型。

[0151] 在信息处理系统100中,以全部客户端模型的分割单位反复执行步骤S109～步骤S110的处理。在全部客户端模型的汇集完成后,进入接下来的步骤S111。另外,在步骤S110中,也可以汇集以分割单位汇集的分割数量的客户端模型并生成1个全局模型。或者,也可以将以分割单位汇集的分割数量的客户端模型设为分割数量的全局模型。

[0152] 最后,服务器装置101的安全执行部105对全局模型进行加密后发送到服务器装置101的通常执行部104。

[0153] 服务器装置101的通常执行部104将被加密的全局模型储存于存储部。

[0154] 具体而言,如下所述。

[0155] 在步骤S111中,服务器装置101的安全执行部105将汇集的客户端模型作为全局模型G,利用发布用的临时秘钥GKs进行加密。服务器装置101的安全执行部105向服务器装置101的通常执行部104发送被加密的全局模型EncGKs(G)。服务器装置101的通常执行部104存储被加密的全局模型EncGKs(G)。

[0156] <全局模型发布>

[0157] 使用图4对信息处理系统100中的全局模型发布处理的动作进行说明。

[0158] 在步骤S112中,服务器装置101的通常执行部104向客户端装置102的通常执行部106发送全局模型的发布通知。或者,也可以从客户端装置102的通常执行部106向服务器装置101的通常执行部104发送全局模型的发布请求。

[0159] 在步骤S113中,服务器装置101的通常执行部104向客户端装置102的通常执行部106发送安全执行部的认证委托,以验证客户端装置102的安全执行部107的正确性。

[0160] 在步骤S114中,客户端装置102的通常执行部106向客户端装置102的安全执行部107发送认证信息的提供委托。

[0161] 在步骤S115中,客户端装置102的安全执行部107向客户端装置102的通常执行部106发送认证信息和公开秘钥PKc。

[0162] 在步骤S116中,客户端装置102的通常执行部106向服务器装置101的通常执行部104转发认证信息和公开秘钥PKc。服务器装置101的通常执行部104向认证服务器装置103的验证部122发送认证信息的验证委托。认证服务器装置103的验证部122向服务器装置101

的通常执行部104发送验证结果。服务器装置101的通常执行部104在能够验证客户端装置102的安全执行部107的正确性的情况下,向服务器装置101的安全执行部105发送公开秘钥PKc。

[0163] 在步骤S117中,服务器装置101的安全执行部105使用公开秘钥PKc与客户端装置102的安全执行部107进行秘钥交换,建立收发数据被加密的安全的通信路径。

[0164] 在步骤S118中,服务器装置101的安全执行部105在安全的通信路径上向客户端装置102的安全执行部107发送发布用的临时秘钥GKs。

[0165] 在步骤S119中,服务器装置101的通常执行部104向客户端装置102的通常执行部106发送被加密的全局模型EncGKs(G)。

[0166] 最后,在步骤S120中,客户端装置102的通常执行部106向客户端装置102的安全执行部107发送被加密的全局模型EncGKs(G),以执行学习或推理处理。客户端装置102的安全执行部107利用发布用的临时秘钥GKs对被加密的全局模型EncGKs(G)进行解密,执行学习或推理处理。

[0167] ***本实施方式的效果的说明***

[0168] 如上所述,根据本实施方式的信息处理系统100,客户端模型和全局模型被加密而在服务器装置101和客户端装置102中交换。此外,客户端模型和全局模型仅在安全执行部105、107中被解密。因此,根据本实施方式的信息处理系统100,能够确保客户端的隐私和全局模型的安全。

[0169] 在本实施方式的信息处理系统100中,服务器装置101和客户端装置102的安全执行部105、107的正确性被验证。因此,根据本实施方式的信息处理系统100,能够防止不正当的服务器装置101和客户端装置102中的不正当的处理。

[0170] 进而,在客户端模型的汇集时检测模型污染,由此,能够防止来自存在恶意的客户端的学习干扰。关于服务器装置101的安全执行部105中的客户端模型的汇集和模型污染检测,通过限制安全执行部105的存储器资源来实现基于分割展开/执行的省存储器化。

[0171] 客户端模型和全局模型在被加密的状态下储存于通常执行部,因此,能够减轻安全执行部的资源负担。

[0172] 不同于客户端模型的加密秘钥,全局模型利用发布用的临时秘钥被加密。由此,模型供应商具有发布用的临时秘钥,还能够调整全局模型。此时,模型供应商不保有客户端模型的加密秘钥,因此,客户端的隐私得到保护。

[0173] ***其他结构***

[0174] 在本实施方式中,服务器装置101、客户端装置102和认证服务器装置103的各装置的功能通过软件来实现。作为变形例,服务器装置101、客户端装置102和认证服务器装置103的各装置的功能也可以通过硬件来实现。

[0175] 具体而言,信息处理系统100代替处理器910而具有电子电路909。

[0176] 图5是示出本实施方式的变形例的信息处理系统100的硬件结构例的图。

[0177] 电子电路909是实现服务器装置101、客户端装置102和认证服务器装置103的各装置的功能的专用的电子电路。具体而言,电子电路909是单一电路、复合电路、程序化的处理器、并行程序化的处理器、逻辑IC、GA、ASIC或FPGA。GA是Gate Array(门阵列)的简称。ASIC是Application Specific Integrated Circuit(专用集成电路)的简称。FPGA是Field-

Programmable Gate Array(现场可编程门阵列)的简称。

[0178] 服务器装置101、客户端装置102和认证服务器装置103的各装置的功能可以通过1个电子电路来实现,也可以分散于多个电子电路来实现。

[0179] 作为另一个变形例,也可以是,服务器装置101、客户端装置102和认证服务器装置103的各装置的一部分的功能通过电子电路来实现,其余的功能通过软件来实现。此外,服务器装置101、客户端装置102和认证服务器装置103的各装置的一部分或全部的功能也可以通过固件来实现。

[0180] 处理器和电子电路分别也被称作处理线路。即,服务器装置101、客户端装置102和认证服务器装置103的各装置的功能通过处理线路来实现。

[0181] 实施方式2

[0182] 在本实施方式中,主要对与实施方式1不同之处和对实施方式1追加之处进行说明。

[0183] 在本实施方式中,对具有与实施方式1相同的功能的结构标注相同标号并省略其说明。

[0184] 在实施方式1中,服务器装置101构成为具有基于TEE的虚拟的分离执行环境。

[0185] 另一方面,在本实施方式中,示出在服务器装置101不具有基于TEE的虚拟的分离执行环境的情况下使用能够在加密的状态下进行运算的同态加密的方式。

[0186] ***结构的说明***

[0187] 图6是示出本实施方式的信息处理系统100的结构例的图。

[0188] 在本实施方式中,服务器装置101仅具有通常的执行环境即通常执行部104。服务器装置101的通常执行部104具有联合学习管理部108、汇集部113和污染检测部114。

[0189] 此外,与实施方式1同样,本实施方式的客户端装置102具有能够虚拟地分离通常执行部106和安全执行部107的结构。

[0190] 客户端装置102的通常执行部106的结构与实施方式1相同。

[0191] 客户端装置102的安全执行部107在与实施方式1相同的结构的基础上具有同态加密解密部140。

[0192] 同态加密解密部140对与服务器装置101交换的模型信息进行同态加密/解密处理。这里,模型信息是客户端模型和全局模型。

[0193] 另外,在本实施方式中,客户端装置102的加密解密部119对与服务器装置101交换的模型信息进行加密/解密处理。

[0194] 与实施方式1同样,认证服务器装置103具有验证部122。在本实施方式中,验证部122验证安全执行部107的认证信息。

[0195] 图6的信息处理系统100通过设为上述这种结构,保护客户端模型和全局模型,验证安全执行部107的正确性。此外,图6的信息处理系统100在服务器装置101的通常执行部104中,在同态加密的状态下进行客户端模型的污染检测和汇集。由此,实现考虑到安全/隐私的联合学习。

[0196] 另外,本实施方式的信息处理系统100的硬件结构例与实施方式1相同。

[0197] ***功能的说明***

[0198] 客户端装置102的安全执行部对向服务器装置101提供的模型信息即客户端模型

执行同态加密。

[0199] 服务器装置101的通常执行部104对被同态加密的客户端模型执行在同态加密的状态下进行汇集的汇集处理。然后,服务器装置101的通常执行部104将通过汇集处理得到的全局模型在被同态加密的状态下储存于存储部。

[0200] 此外,服务器装置101的通常执行部104对被同态加密的客户端模型执行在同态加密的状态下检测污染的污染检测处理。

[0201] ***功能的详细说明***

[0202] 接着,使用图6更加详细地说明信息处理系统100的各装置的功能。有时省略实施方式1中说明的部分。

[0203] 以联合学习为代表的分布式机器学习算法通过服务器装置101和客户端装置102各自的联合学习管理部108、115彼此的交换123来执行。设想客户端装置102存在多个。

[0204] 客户端装置102的通常执行部106和安全执行部107的虚拟分离例如通过Arm Trustzone或Intel(注册商标)SGX这样的TEE技术来实现。

[0205] 联合学习管理部108、115进行联合学习用的客户端模型的收集或全局模型的发布。此外,客户端装置102的联合学习管理部115通过认证管理部116来验证安全执行部107的正确性(处理125)。

[0206] 对用于验证安全执行部107的正确性的结构要素进行说明。

[0207] 认证管理部116从位于安全执行部107的认证部118取得用于验证安全执行部107的正确性的认证信息(处理127)。

[0208] 认证部118输出认证信息(处理127)。认证信息例如是已启动的安全执行部的哈希值和签名。安全执行部107的认证例如通过Remote Attestation技术来实现。

[0209] 验证部122取得来自认证管理部116的认证信息,验证安全执行部107是否正确地启动(处理129)。

[0210] 对服务器装置101具有的结构要素进行说明。

[0211] 汇集部113取得由联合学习管理部108收集到的被同态加密的客户端模型(处理225),进行汇集。汇集例如是指计算客户端模型的平均值。但是,成为被同态加密的状态下的运算。

[0212] 污染检测部114取得由联合学习管理部108收集到的被同态加密的客户端模型(处理226),进行客户端模型的污染检测。污染检测例如是指,计算客户端模型之间的模型间距离,在距离大的情况下,检测为该客户端模型被污染。但是,成为被同态加密的状态下的运算,因此,在客户端装置102中进行距离大小的判定。

[0213] 对客户端装置102具有的结构要素进行说明。

[0214] 学习/推理管理部117使用从服务器装置101发布的全局模型对学习或推理处理的执行进行管理(处理136)。

[0215] 同态加密解密部140对通过联合学习管理部115向服务器装置101提供的客户端模型进行同态加密处理(处理223)。或者,同态加密解密部140对从服务器装置101发布的被同态加密的全局模型进行解密处理(处理224)。

[0216] 加密解密部119对同态加密被解密的全局模型进行重加密。或者,加密解密部119对被加密的模型信息进行解密(处理223)。

[0217] 学习部120通过学习/推理管理部117,使用由加密解密部119解密后的全局模型(处理137)来执行学习。

[0218] 推理部121通过学习/推理管理部117,使用由加密解密部119解密后的全局模型(处理138)来执行推理。

[0219] ***动作的说明***

[0220] 接着,对本实施方式的信息处理系统100的动作进行说明。信息处理系统100的动作步骤相当于信息处理方法。此外,实现信息处理系统100的动作的程序相当于信息处理程序。

[0221] 图7是示出本实施方式的信息处理系统100中的客户端模型收集的动作的序列图。

[0222] 图8是示出本实施方式的信息处理系统100中的全局模型发布的动作的序列图。

[0223] 该序列图分成通常执行部104、106和安全执行部107来示出本实施方式的信息处理系统100中的服务器装置101和客户端装置102的交换。

[0224] <客户端模型收集>

[0225] 使用图7对本实施方式的信息处理系统100中的客户端模型收集处理的动作进行说明。

[0226] 在步骤S201中,服务器装置101的通常执行部104向客户端装置102的通常执行部106发送客户端模型的提供委托。

[0227] 在步骤S202中,客户端装置102的通常执行部106从客户端装置102的安全执行部107取得被同态加密的客户端模型HEMKc(M)。

[0228] 在步骤S203中,客户端装置102的通常执行部106向服务器装置101的通常执行部104发送被同态加密的客户端模型HEMKc(M)。

[0229] 在各客户端中执行以上的步骤S201～步骤S203,服务器装置101收集客户端模型。在全部客户端模型的收集完成后,执行接下来的步骤S204。

[0230] 最后,在步骤S204中,服务器装置101的通常执行部104使用被同态加密的客户端模型HEMKc(M),在加密的状态下执行污染检测和汇集。服务器装置101的通常执行部104将被汇集的客户端模型作为全局模型,将被同态加密的全局模型HEGKs(G)和污染检测结果储存于存储部。

[0231] <全局模型发布处理>

[0232] 使用图8对本实施方式的信息处理系统100中的全局模型发布处理的动作进行说明。

[0233] 在步骤S205中,服务器装置101的通常执行部104向客户端装置102的通常执行部106发送全局模型的发布通知。也可以从客户端装置102的通常执行部106向服务器装置101的通常执行部104发送全局模型的发布请求。

[0234] 在步骤S206中,服务器装置101的通常执行部104向客户端装置102的通常执行部106发送安全执行部的认证委托,以验证客户端装置102的安全执行部107的正确性。

[0235] 在步骤S207中,客户端装置102的通常执行部106向客户端装置102的安全执行部107发送认证信息的提供委托。

[0236] 在步骤S208中,客户端装置102的安全执行部107向客户端装置102的通常执行部106发送认证信息和公开秘钥PKc。

[0237] 在步骤S209中,客户端装置102的通常执行部106向服务器装置101的通常执行部104转发认证信息和公开秘钥PKc。服务器装置101的通常执行部104向认证服务器装置103的验证部122发送认证信息的验证委托。认证服务器装置103的验证部122向服务器装置101的通常执行部104发送验证结果。服务器装置101的通常执行部104在能够验证客户端装置102的安全执行部107的正确性的情况下,向服务器装置101的通常执行部104发送公开秘钥PKc。

[0238] 在步骤S210中,服务器装置101的通常执行部104使用公开秘钥PKc与客户端装置102的安全执行部107进行秘钥交换,建立收发数据被加密的安全的通信路径。

[0239] 在步骤S211中,服务器装置101的通常执行部104在安全的通信路径上向客户端装置102的安全执行部107发送被同态加密的全局模型HEGKs(G)和污染检测结果。

[0240] 在步骤S212中,客户端装置102的安全执行部107对被同态加密的全局模型HEGKs(G)和污染检测结果进行解密。客户端装置102的安全执行部107根据污染检测结果,如果客户端模型没有污染,则利用客户端的模型保护用秘钥GKc对全局模型进行加密。然后,客户端装置102的安全执行部107向客户端装置102的通常执行部106发送被加密的全局模型EncGKc(G)。

[0241] 最后,在步骤S214中,客户端装置102的通常执行部106向客户端装置102的安全执行部107发送被加密的全局模型EncGKc(G),以执行学习或推理处理。客户端装置102的安全执行部107利用客户端的模型保护用秘钥GKc对被加密的全局模型EncGKc(G)进行解密,执行学习或推理处理。

[0242] ***本实施方式的效果的说明***

[0243] 如上所述,根据本实施方式的信息处理系统100,客户端模型和全局模型被同态加密而在服务器装置101和客户端装置102中交换。而且,客户端模型和全局模型在通过同态加密进行加密的状态下被运算,或者仅在客户端装置102的安全执行部中被解密。因此,能够确保客户端的隐私和全局模型的安全。

[0244] 此外,客户端装置102的安全执行部的正确性得到验证。因此,能够防止不正当的客户端装置102中的不正当的处理。进而,在客户端装置102的安全执行部中确认全局模型和污染检测结果,由此,能够防止来自存在恶意的客户端的学习干扰。

[0245] 客户端模型和全局模型在被加密的状态下储存于通常执行部,因此,能够减轻安全执行部的资源负担。

[0246] 实施方式3

[0247] 在本实施方式中,主要对与实施方式1、2不同之处和对实施方式1、2追加之处进行说明。

[0248] 在本实施方式中,对具有与实施方式1、2相同的功能的结构标注相同标号并省略其说明。

[0249] 在实施方式2中,说明了在服务器装置101不具有基于TEE的虚拟的分离执行环境的情况下使用能够在加密的状态下进行运算的同态加密的方式。

[0250] 另一方面,在本实施方式中,说明如下方式:在服务器装置101中,一并使用基于TEE的虚拟的分离执行环境和能够在加密的状态下进行运算的同态加密。

[0251] ***结构的说明***

[0252] 图9是示出本实施方式的信息处理系统100的结构例的图。

[0253] 与实施方式1同样,服务器装置101具有能够虚拟地分离成通常执行部104和安全执行部105的结构。

[0254] 此外,与实施方式1、2同样,客户端装置102也具有能够虚拟地分离成通常执行部106和安全执行部107的结构。

[0255] 服务器装置101的通常执行部104具有联合学习管理部108、认证管理部109、汇集部113和污染检测部114。

[0256] 服务器装置101的安全执行部105具有认证部110、加密解密部111和同态加密解密部140。在本实施方式中,同态加密解密部140对与服务器装置101的通常执行部104交换的信息进行同态加密/解密处理。

[0257] 与实施方式1同样,客户端装置102的通常执行部104具有联合学习管理部115、认证管理部116和学习/推理管理部117。

[0258] 与实施方式1同样,客户端装置102的安全执行部107具有认证部118、加密解密部119、学习部120和推理部121。

[0259] 认证服务器装置103具有验证部122。

[0260] 验证部122验证安全执行部105和安全执行部107各自的认证信息。

[0261] 图9的信息处理系统100通过设为上述这种结构,保护客户端模型和全局模型,检测安全执行部105、107各自的正确性和客户端模型的污染。由此,实现考虑到安全/隐私的联合学习。

[0262] 另外,本实施方式的信息处理系统100的硬件结构例与实施方式1相同。

[0263] ***功能的说明***

[0264] 通常执行部104、106彼此认证安全执行部105、107的启动的正确性。在安全执行部105、107的启动的正确性得到认证时,在安全执行部105、107彼此之间建立收发被加密的数据的安全的通信路径。

[0265] 服务器装置101的安全执行部105对从客户端装置102经由安全的通信路径提供的模型信息即客户端模型执行同态加密。然后,服务器装置101的安全执行部105将被同态加密的模型信息在被同态加密的状态下储存于存储部。

[0266] 服务器装置101的通常执行部104对被同态加密的模型信息执行在同态加密的状态下进行汇集的汇集处理。然后,服务器装置101的通常执行部104将通过汇集处理得到的模型信息即全局模型在被同态加密的状态下储存于存储部。

[0267] 此外,服务器装置101的通常执行部104对被同态加密的模型信息即客户端模型执行在同态加密的状态下检测污染的污染检测处理。

[0268] ***功能的详细说明***

[0269] 接着,使用图9更加详细地说明信息处理系统100的各装置的功能。

[0270] 以联合学习为代表的分布式机器学习算法通过服务器装置101和客户端装置102各自的联合学习管理部108、115彼此的交换123来执行。设想客户端装置102存在多个。

[0271] 客户端装置102的通常执行部106和安全执行部107的虚拟分离例如通过Arm Trustzone或Intel(注册商标)SGX这样的TEE技术来实现。

[0272] 联合学习管理部108、115通过各认证管理部109、116来验证彼此的安全执行部

105、107的正确性。该处理与实施方式1中说明的处理相同。

[0273] 加密解密部111按照每个客户端,对通过联合学习管理部108从客户端装置102收集到的客户端模型进行解密处理。或者,加密解密部111按照每个客户端,对通过联合学习管理部108向客户端装置102发布的全局模型进行加密处理(处理130)。

[0274] 同态加密解密部140利用临时的公共秘钥对收集到的客户端模型进行同态加密处理(处理331)后储存于通常执行部104。或者,同态加密解密部140从通常执行部104取得被同态加密的状态的全局模型,进行解密处理(处理332)。

[0275] 汇集部113取得收集到的已同态加密的客户端模型(处理332),进行汇集。汇集例如是指计算客户端模型的平均值。但是,在本实施方式中,成为被同态加密的状态下的运算。

[0276] 污染检测部114取得收集到的已同态加密的客户端模型(处理333),在同态加密的状态下进行客户端模型的污染检测。污染检测例如是指,计算客户端模型之间的模型间距离,在距离大的情况下,检测为该客户端模型被污染。但是,在本实施方式中,成为被同态加密的状态下的运算,因此,在服务器装置101的安全执行部105中进行解密后,进行距离大小的判定。

[0277] 客户端装置102中的向服务器装置101提供客户端模型的功能、以及使用从服务器装置101发布的全局模型进行学习或推理的功能与实施方式1中说明的情况相同。

[0278] ***动作的说明***

[0279] 接着,对本实施方式的信息处理系统100的动作进行说明。信息处理系统100的动作步骤相当于信息处理方法。此外,实现信息处理系统100的动作的程序相当于信息处理程序。

[0280] 图10是示出本实施方式的信息处理系统100中的客户端模型收集的动作的序列图。

[0281] 图11是示出本实施方式的信息处理系统100中的全局模型发布的动作的序列图。

[0282] 该序列图分成通常执行部104、106以及安全执行部105、107来示出信息处理系统100中的服务器装置101和客户端装置102的交换。

[0283] <客户端模型收集>

[0284] 步骤S301～步骤S307的处理与实施方式1中说明的步骤S101～步骤S107的处理相同。即,在步骤S307中,客户端装置102的安全执行部107在安全的通信路径上向服务器装置101的安全执行部105发送客户端模型M。

[0285] 在步骤S308中,服务器装置101的安全执行部105一次性地利用运算用的临时秘钥MKs对客户端模型M进行同态加密,以抑制安全执行部105的消耗存储器。然后,服务器装置101的安全执行部105向服务器装置101的通常执行部104发送被同态加密的客户端模型HEMKs(M)。服务器装置101的通常执行部104存储被同态加密的客户端模型HEMKs(M)。

[0286] 在各客户端中执行以上的步骤S301～步骤S308,从全部客户端装置102收集客户端模型。在全部客户端模型的收集完成后,执行接下来的步骤S309。

[0287] 在步骤S309中,服务器装置101的通常执行部104使用被同态加密的客户端模型HEMKs(M),在加密的状态下执行污染检测和汇集。

[0288] 在步骤S310中,服务器装置101的通常执行部104将被汇集的客户端模型作为全局

模型,向服务器装置101的安全执行部105发送被同态加密的全局模型HEGKs(G)和污染检测结果。

[0289] 最后,在步骤S311中,服务器装置101的安全执行部105对被同态加密的全局模型HEGKs(G)和污染检测结果进行解密。在检测到污染的情况下,不汇集被污染的客户端模型。例如,在检测到被污染的客户端模型的情况下,也可以丢弃全局模型。服务器装置101的安全执行部105利用发布用的临时秘钥GKs对全局模型G进行加密,向服务器装置101的通常执行部104发送被加密的全局模型EncGKs(G)。服务器装置101的通常执行部104存储被加密的全局模型EncGKs(G)。

[0290] <全局模型发布>

[0291] 使用图11对信息处理系统100中的全局模型发布处理的动作进行说明。

[0292] 步骤S312～步骤S320的处理与实施方式1中说明的步骤S112～步骤S120的处理相同。

[0293] 即,最后,在步骤S320中,客户端装置102的通常执行部106向客户端装置102的安全执行部107发送被加密的全局模型EncGKs(G),以执行学习或推理处理。客户端装置102的安全执行部107利用发布用的临时秘钥GKs对被加密的全局模型EncGKs(G)进行解密,执行学习或推理处理。

[0294] ***本实施方式的效果的说明***

[0295] 如上所述,在本实施方式的信息处理系统100中,客户端模型和全局模型被加密而在服务器装置101和客户端装置102中交换。此外,在服务器装置101的通常执行部中,在通过同态加密进行加密的状态下被运算。此外,仅在服务器装置101和客户端装置102的各装置的安全执行部中被解密。因此,能够确保客户端的隐私和全局模型的安全。

[0296] 此外,在本实施方式的信息处理系统100中,服务器装置101和客户端装置102的各安全执行部的正确性得到验证。因此,能够防止不正当的服务器装置101和客户端装置102中的不正当的处理。

[0297] 进而,在汇集客户端模型时检测模型污染,由此,能够防止来自存在恶意的客户端的学习干扰。

[0298] 此外,通过限制安全执行部的存储器资源,在使用同态加密进行加密的状态下,在具有充裕的存储器/计算资源的通常执行部中进行运算,由此,实现服务器装置101中的客户端模型的汇集和模型污染检测。

[0299] 不同于客户端模型的加密秘钥,全局模型利用发布用的临时秘钥进行加密。由此,模型供应商具有发布用的临时秘钥,还能够调整全局模型。此时,模型供应商不保有客户端模型的加密秘钥,因此,客户端的隐私得到保护。

[0300] 实施方式4

[0301] 在本实施方式中,主要对与实施方式1不同之处和对实施方式1追加之处进行说明。

[0302] 在本实施方式中,对具有与实施方式1相同的功能的结构标注相同标号并省略其说明。

[0303] 在实施方式1中,构成为服务器装置101的安全执行部105具有污染检测部114。在本实施方式中,对客户端装置102的安全执行部107具有污染检测部114的方式进行说明。

[0304] ***结构的说明***

[0305] 图12是示出本实施方式的信息处理系统100的结构例的图。

[0306] 在本实施方式中,在实施方式1中说明的服务器装置101的安全执行部105中没有污染检测部114。在实施方式1中说明的客户端装置102的安全执行部107中具有污染检测部114。

[0307] 客户端装置102的安全执行部107执行检测向服务器装置101提供的客户端模型是否被污染的污染检测处理。然后,客户端装置102的安全执行部107不向服务器装置101提供被污染的客户端模型。

[0308] 除了上述以外,与实施方式1中说明的信息处理系统100的结构相同。

[0309] 在客户端装置102中,在安全执行部107中具有认证部118、加密解密部119、污染检测部114、学习部120和推理部121。

[0310] 污染检测部114检测向服务器装置101提供的客户端模型的污染。

[0311] 在图12的信息处理系统100中,通过设为上述这种结构,保护客户端模型和全局模型,检测安全执行部105、107各自的正确性和客户端模型的污染。由此,实现考虑到安全/隐私的联合学习。

[0312] ***功能的详细说明***

[0313] 接着,使用图12更加详细地说明信息处理系统100的各装置的功能。

[0314] 以联合学习为代表的分布式机器学习算法通过服务器装置101和客户端装置102各自的联合学习管理部108、115彼此的交换123来执行。设想客户端装置102存在多个。

[0315] 联合学习管理部108、115通过各认证管理部109、116验证彼此的安全执行部105、107的正确性。该处理与实施方式1中说明的处理相同。

[0316] 服务器装置101中的加密解密部111进行的处理、重加密解密部112进行的处理和汇集部113进行的处理也与实施方式1相同。

[0317] 客户端装置102中的加密解密部119进行的处理也与实施方式1相同。

[0318] 客户端装置102的污染检测部114进行向服务器装置101提供的客户端模型的污染检测(处理435)。污染检测例如是指,计算客户端模型与原来的全局之间的模型间距离,在距离大的情况下检测为该客户端模型被污染,或者,根据针对特定的测试数据的输出结果检测被污染。

[0319] 客户端装置102中的学习/推理管理部117进行的处理、学习部120进行的处理和推理部121进行的处理与实施方式1相同。但是,在学习/推理管理部117进行的处理、学习部120进行的处理和推理部121进行的处理中不使用被检测到污染的客户端模型。

[0320] 另外,本实施方式的信息处理系统100的硬件结构例与实施方式1相同。

[0321] ***动作的说明***

[0322] 关于动作,在向服务器装置101提供客户端模型之前,通过客户端装置102的安全执行部107中的污染检测部114实施实施方式1中的污染检测部114的处理。其他没有变更。

[0323] 此外,也可以对实施方式2、3应用本实施方式。在实施方式2、3中,也可以在客户端装置102的安全执行部107中具有污染检测部114。客户端装置102的安全执行部107执行检测向服务器装置101提供的客户端模型是否被污染的污染检测处理。然后,客户端装置102的安全执行部107不向服务器装置101提供被污染的客户端模型。

[0324] 在以上的实施方式1～4中,将信息处理系统的各装置的各部作为独立的功能块进行了说明。但是,信息处理系统的各装置的结构也可以不是上述的实施方式这样的结构。信息处理系统的各装置的功能块能够实现上述的实施方式中说明的功能即可,也可以是任意的结构。此外,信息处理系统的各装置可以是1个装置,也可以是由多个装置构成的系统。

[0325] 此外,也可以组合实施实施方式1～4中的多个部分。或者,也可以实施这些实施方式中的1个部分。除此之外,也可以整体或部分地任意组合实施这些实施方式。

[0326] 即,在实施方式1～4中,能够进行各实施方式的自由组合、或各实施方式的任意结构要素的变形、或各实施方式中的任意结构要素的省略。

[0327] 另外,上述的实施方式是本质上优选的例示,并不意图限制本发明的范围、本发明的应用物的范围和本发明的用途的范围。上述的实施方式能够根据需要进行各种变更。

[0328] 标号说明

[0329] 100:信息处理系统;101:服务器装置;102:客户端装置;103:认证服务器装置;104、106:通常执行部;105、107:安全执行部;108、115:联合学习管理部;109、116:认证管理部;110、118:认证部;111、119:加密解密部;112:重加密解密部;113:汇集部;114:污染检测部;117:学习/推理管理部;120:学习部;121:推理部;122:验证部;140:同态加密解密部;909:电子电路;910:处理器;921:存储器;922:辅助存储装置;930:输入接口;940:输出接口;950:通信装置。

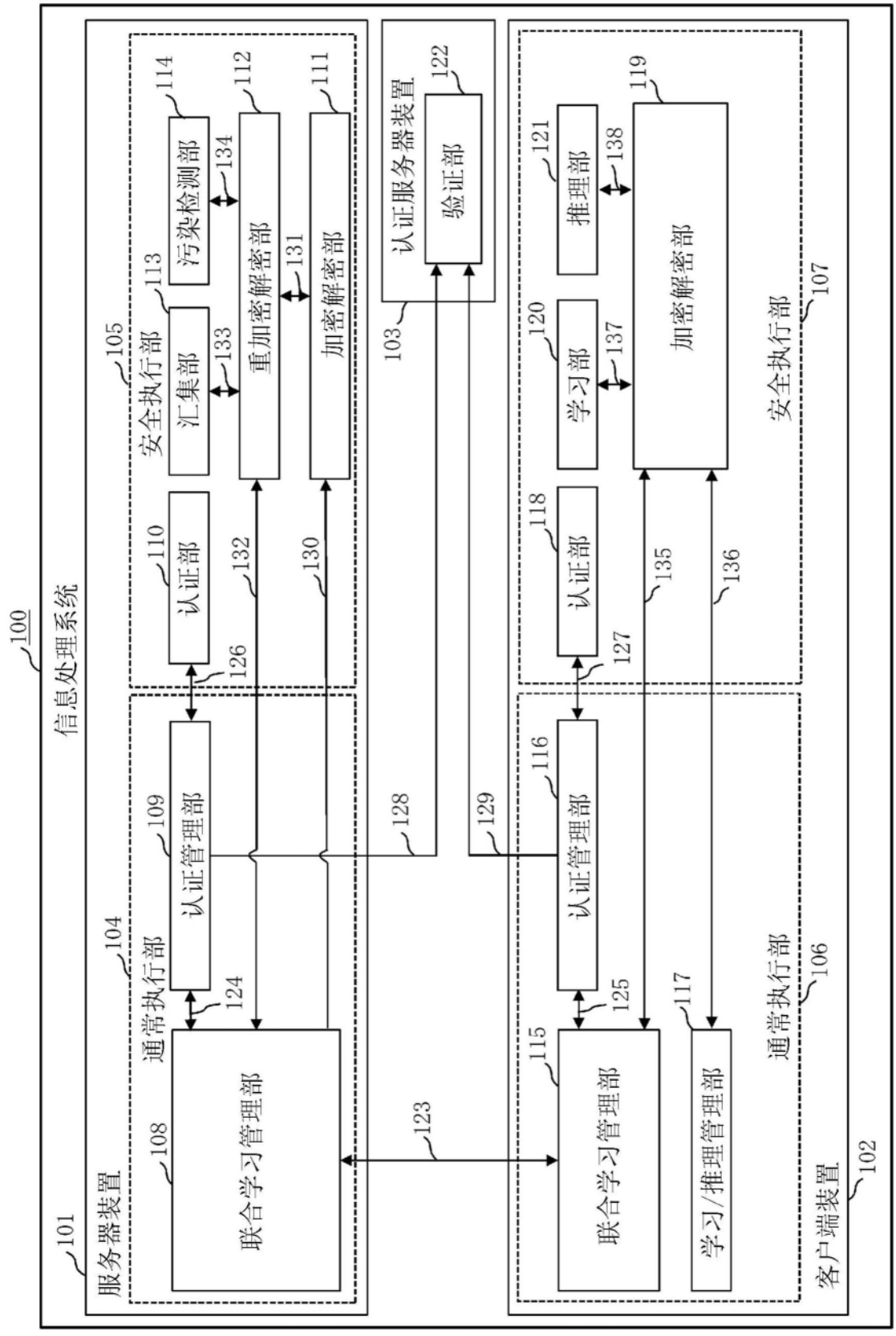

100
信息处理系统
101
服务器装置
104
通常执行部
108
联合学习管理部
109
认证管理部
124
105
安全执行部
110
认证部
126
113
汇集部
133
114
污染检测部
134
112
重加密解密部
131
111
加密解密部
132
130
103
认证服务器装置
122
验证部
128
129
123
102
客户端装置
106
通常执行部
115
联合学习管理部
116
认证管理部
125
117
学习/推理管理部
107
安全执行部
118
认证部
127
120
学习部
137
121
推理部
138
119
加密解密部
135
136

图1

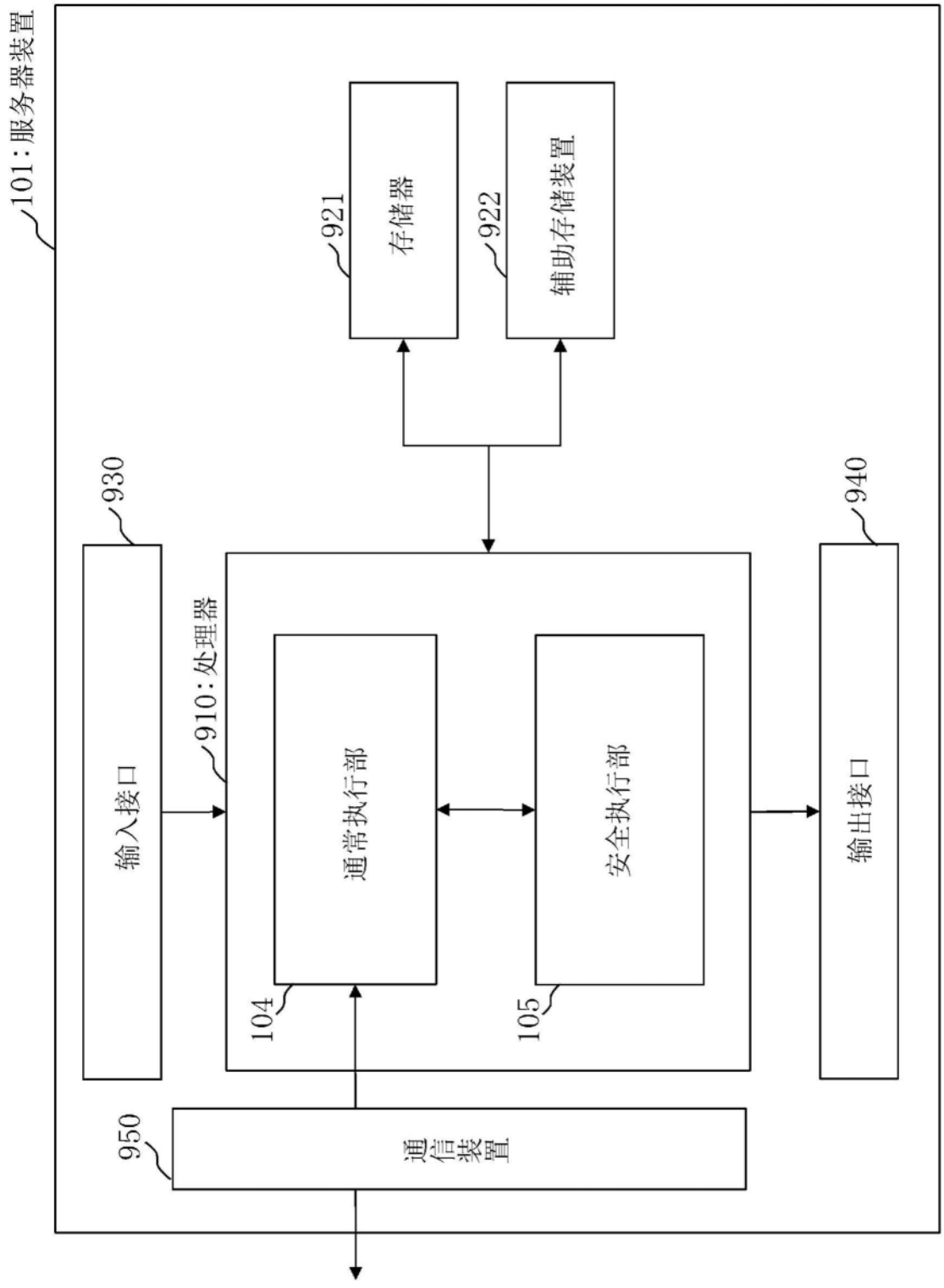

101:服务器装置
921
存储器
922
辅助存储装置
930
940
输入接口
910:处理器
通常执行部
安全执行部
输出接口
104
105
950
通信装置

图2

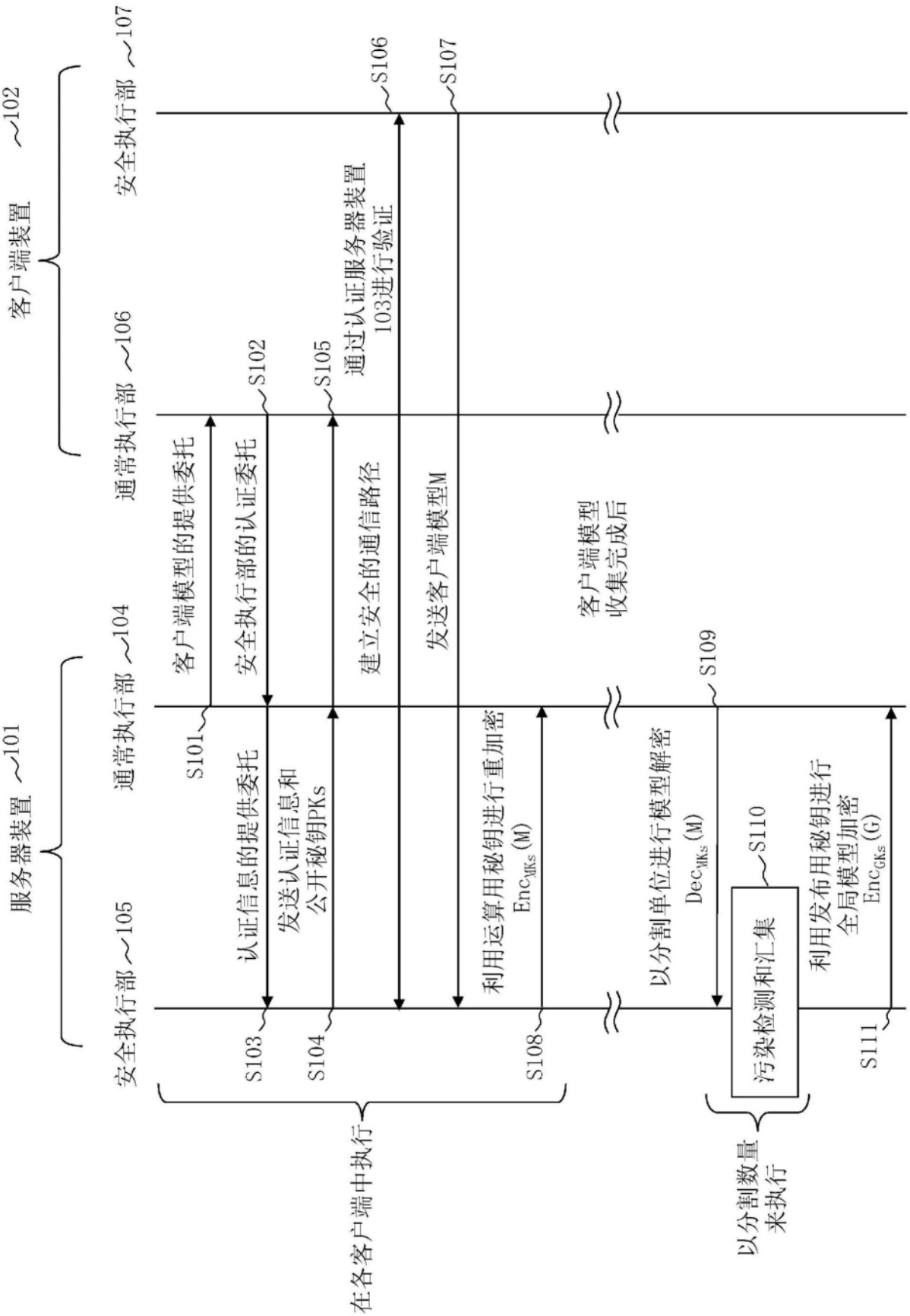

服务器装置 101
客户端装置 102
安全执行部 105
通常执行部 104
通常执行部 106
安全执行部 107
S101
客户端模型的提供委托
S102
安全执行部的认证委托
S103
认证信息的提供委托
S104
发送认证信息和
公开秘钥PKs
S105
建立安全的通信路径
S106
通过认证服务器装置
103进行验证
S107
发送客户端模型M
S108
利用运算用秘钥进行重加密
Enc_{VKs} (M)
客户端模型
收集完成后
S109
以分割单位进行模型解密
Dec_{VKs} (M)
S110
污染检测和汇集
利用发布用秘钥进行
全局模型加密
Enc_{GKs} (G)
S111
在各客户端中执行
以分割数量
来执行

图3

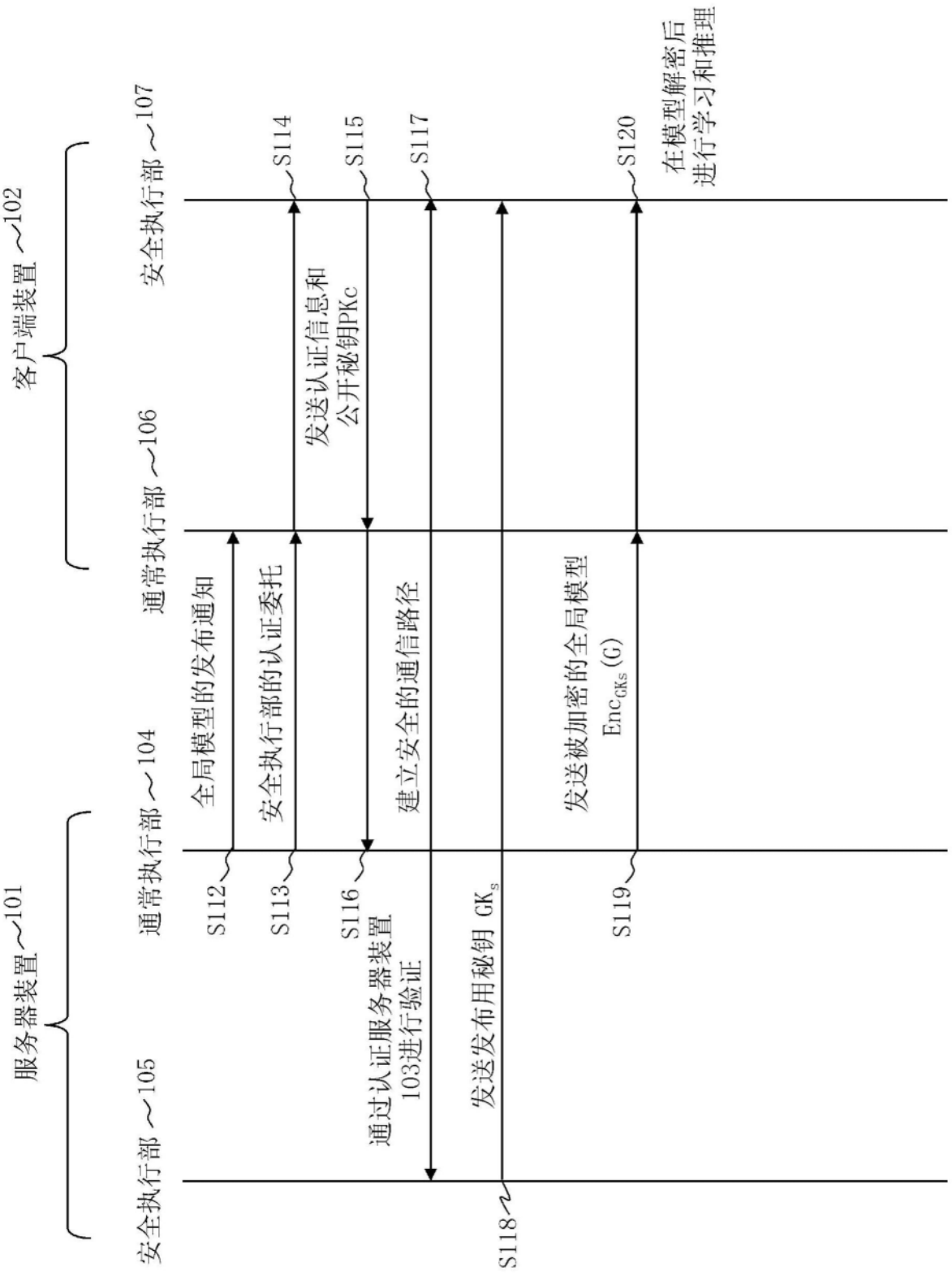

客户端装置 ~102
服务器装置 ~101
安全执行部 ~105
通常执行部 ~104
通常执行部 ~106
安全执行部 ~107
全局模型的发布通知
S112
安全执行部的认证委托
S113
通过认证服务器装置
103进行验证
S116
发送发布用秘钥 GK_S
S118
发送认证信息和
公开秘钥PKc
S114
S115
建立安全的通信路径
S117
发送被加密的全局模型
$Enc_{GK_S}(G)$
S119
S120
在模型解密后
进行学习和推理

图4

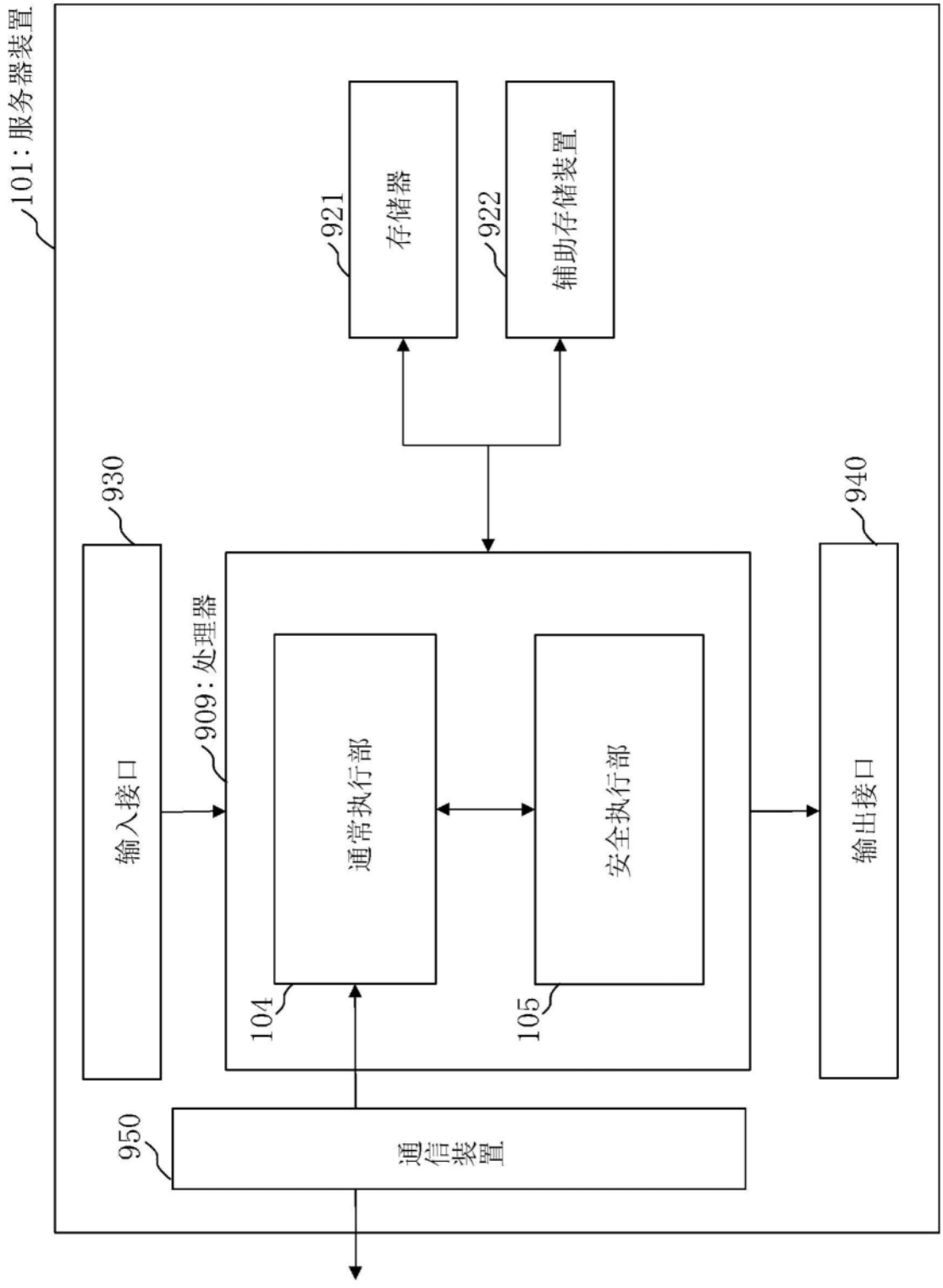

101：服务器装置
921
存储器
922
辅助存储装置
930
940
输入接口
909：处理器
通常执行部
安全执行部
输出接口
104
105
950
通信装置

图5

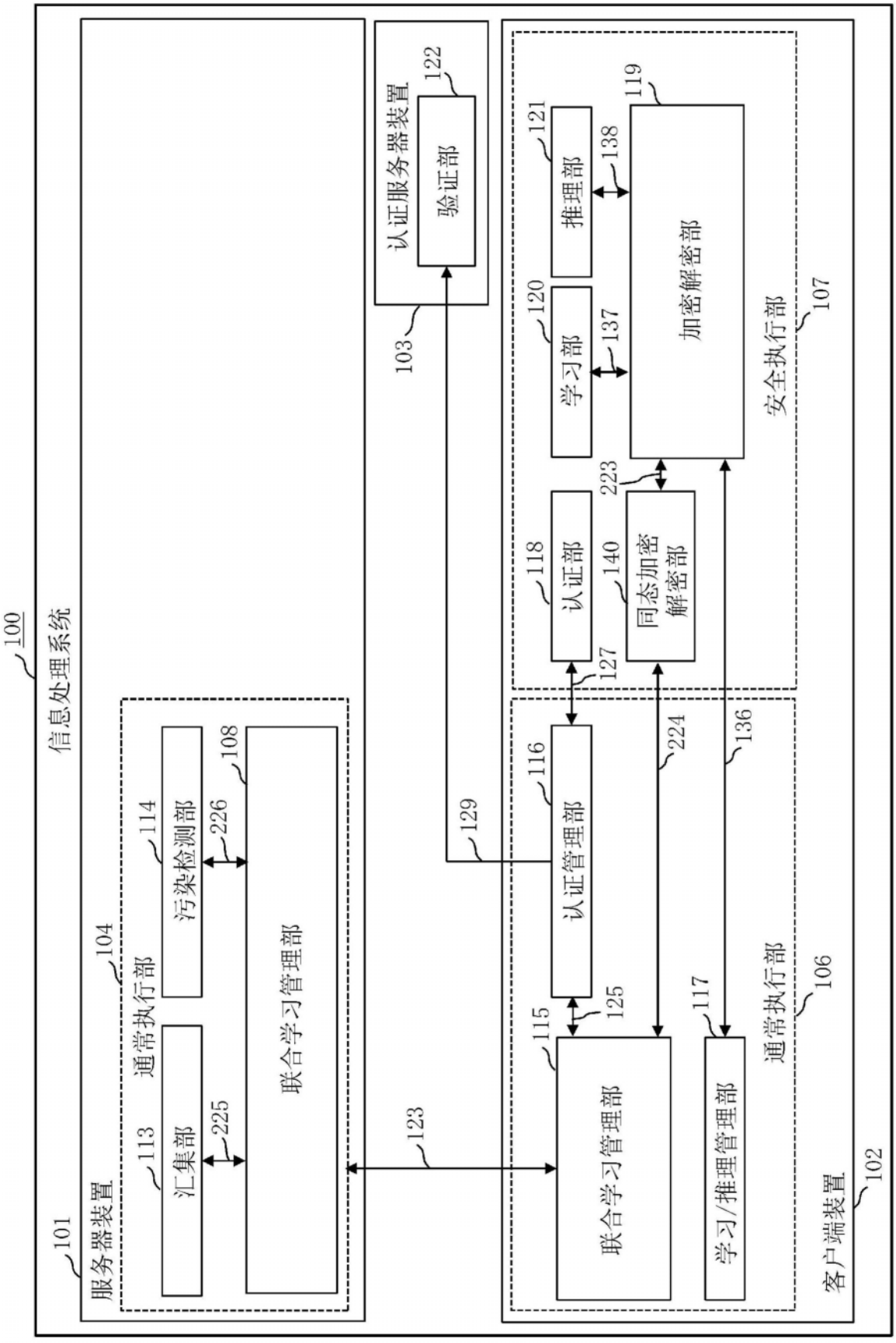

100
信息处理系统
101
服务器装置
104
通常执行部
113
汇集部
114
污染检测部
108
联合学习管理部
225
226
123
103
认证服务器装置
122
验证部
129
102
客户端装置
106
通常执行部
115
联合学习管理部
125
116
认证管理部
117
学习/推理管理部
136
224
107
安全执行部
118
认证部
127
140
同态加密解密部
223
120
学习部
137
121
推理部
138
119
加密解密部

图6

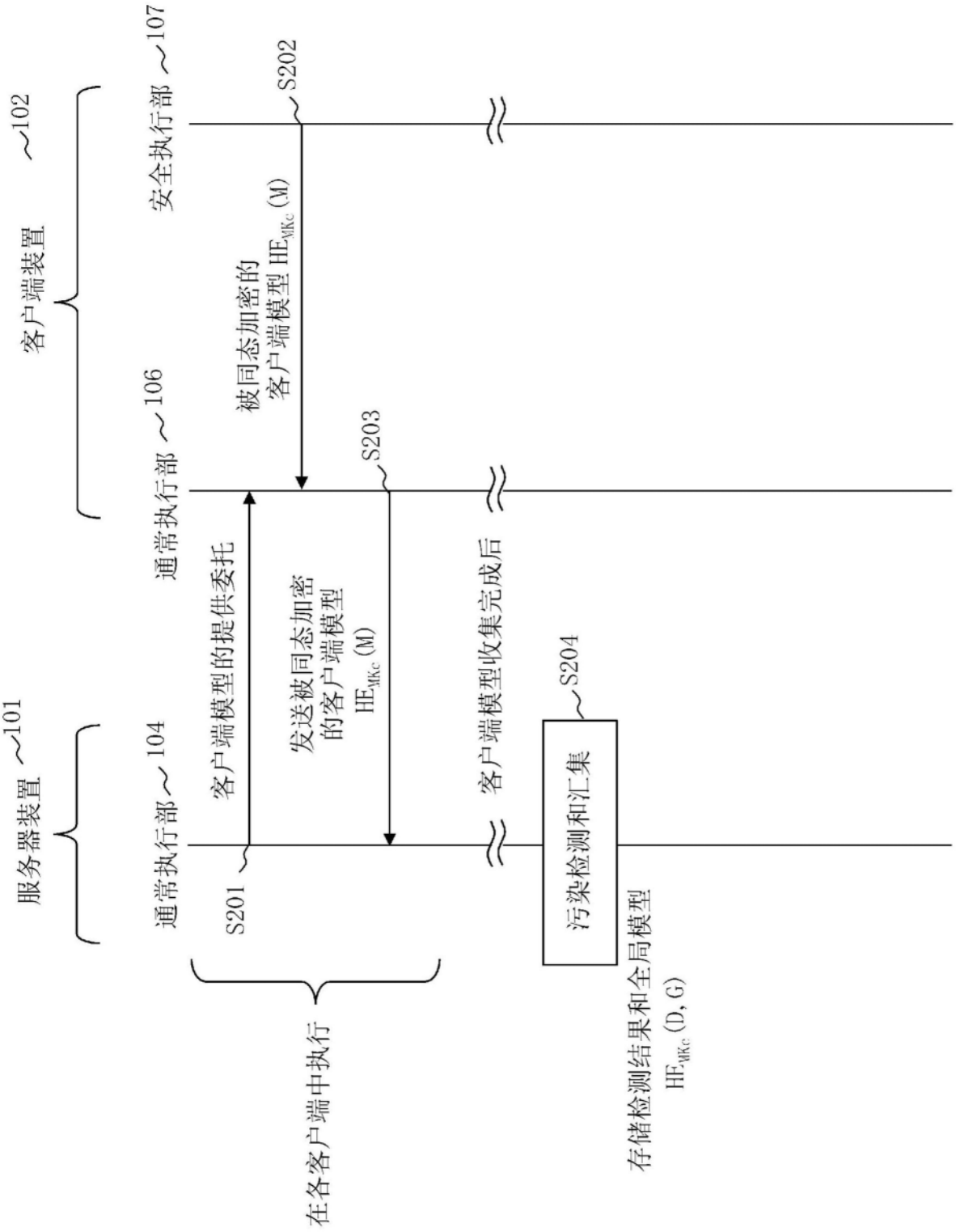

服务器装置
101
客户端装置
102
通常执行部
104
通常执行部
106
安全执行部
107
客户端模型的提供委托
S201
发送被同态加密的客户端模型 $HE_{MKc}(M)$
被同态加密的客户端模型 $HE_{MKc}(M)$
S202
S203
客户端模型收集完成后
S204
污染检测和汇集
在各客户端中执行
存储检测结果和全局模型 $HE_{MKc}(D,G)$

图7

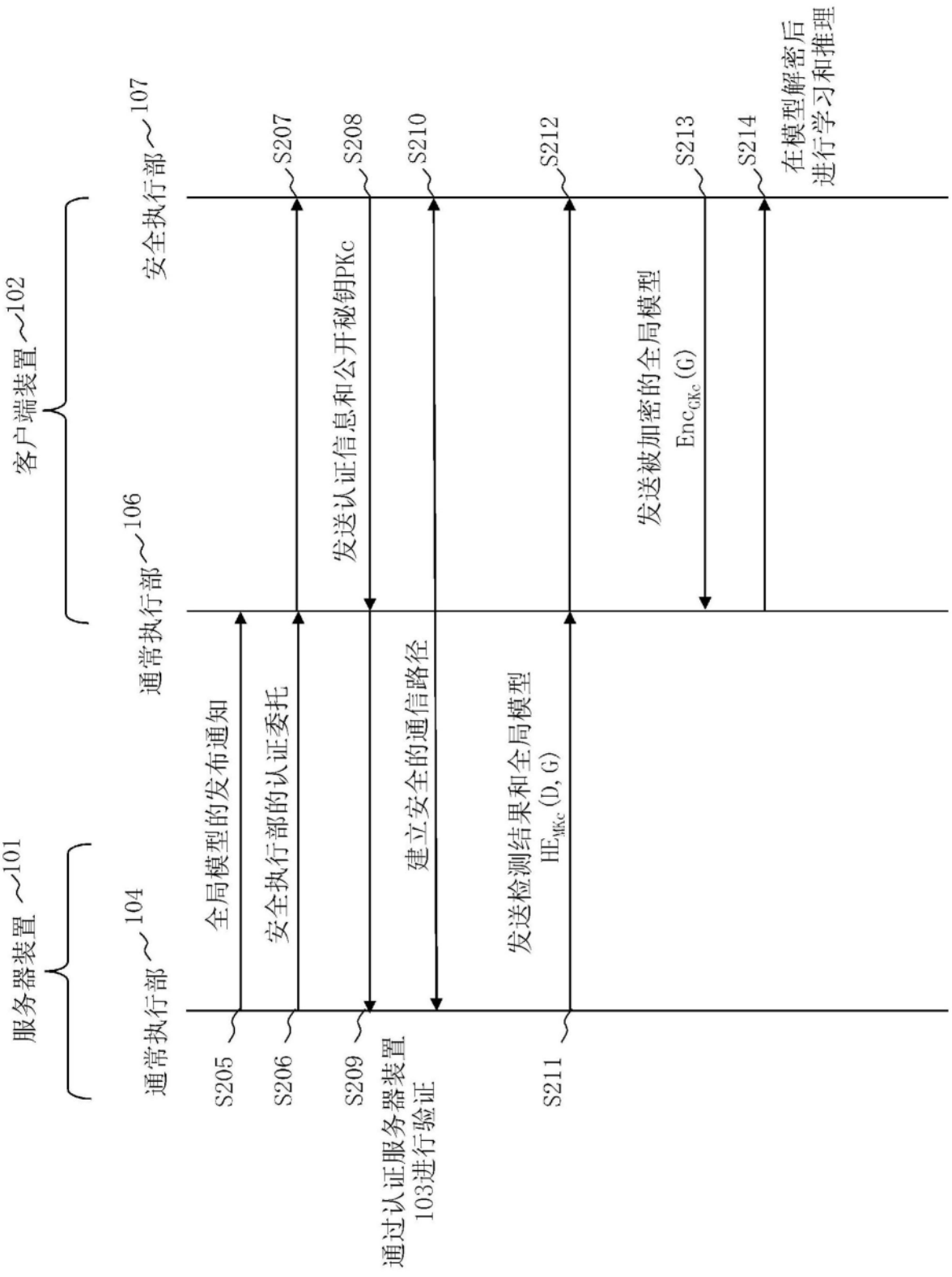

服务器装置 101
客户端装置 102
通常执行部 104
通常执行部 106
安全执行部 107
S205
全局模型的发布通知
S206
安全执行部的认证委托
S207
S208
发送认证信息和公开秘钥PKc
S209
通过认证服务器装置103进行验证
建立安全的通信路径
S210
S211
发送检测结果和全局模型
$HE_{GKc}(D, G)$
S212
S213
发送被加密的全局模型
$Enc_{GKc}(G)$
S214
在模型解密后进行学习和推理

图8

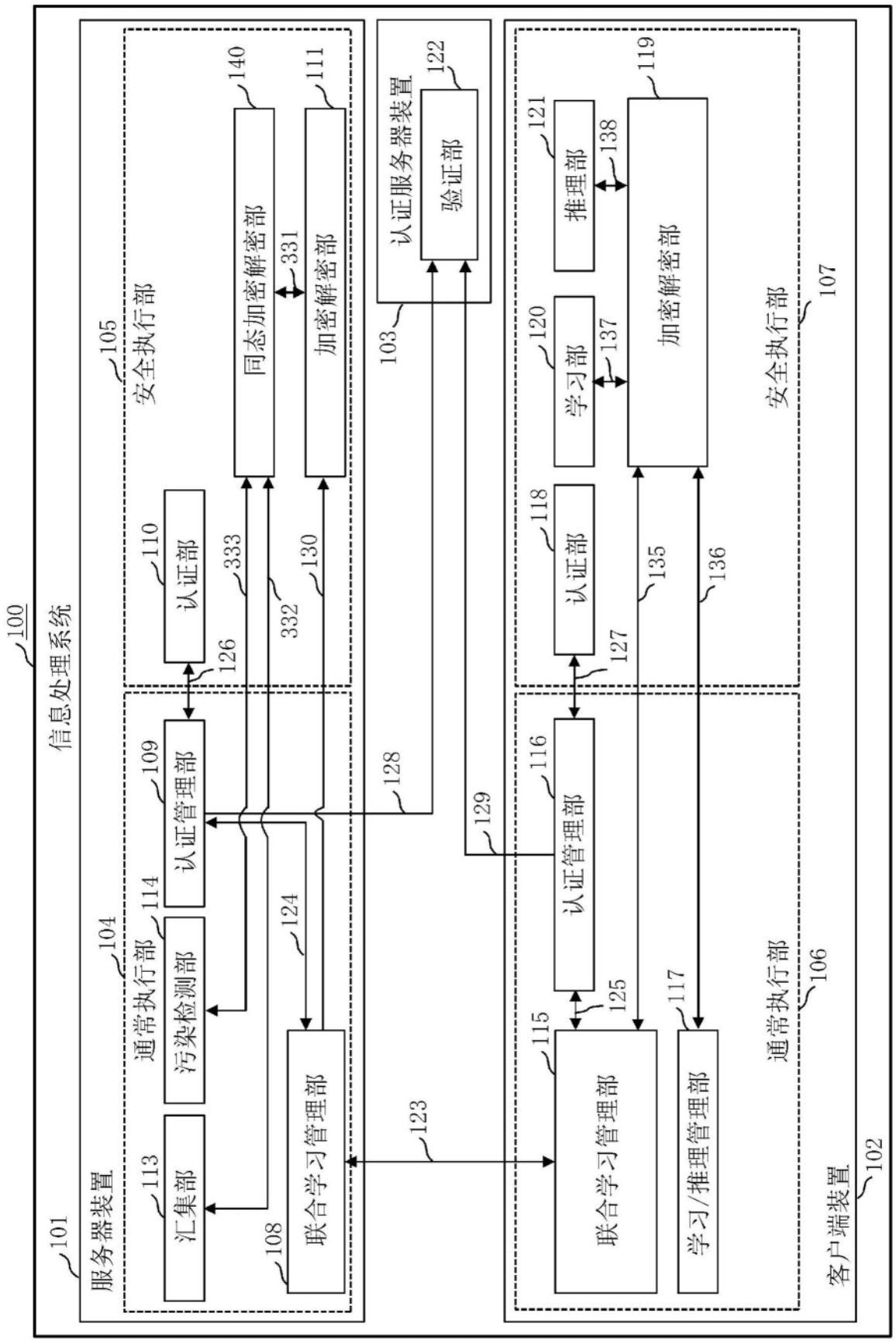

100
信息处理系统
101
服务器装置
105
安全执行部
110
认证部
126
140
同态加密解密部
331
111
加密解密部
333
332
130
104
通常执行部
114
污染检测部
109
认证管理部
113
汇集部
108
联合学习管理部
124
128
129
123
103
认证服务器装置
122
验证部
107
安全执行部
121
推理部
138
119
加密解密部
120
学习部
137
118
认证部
135
136
127
116
认证管理部
125
115
联合学习管理部
117
学习/推理管理部
106
通常执行部
102
客户端装置

图9

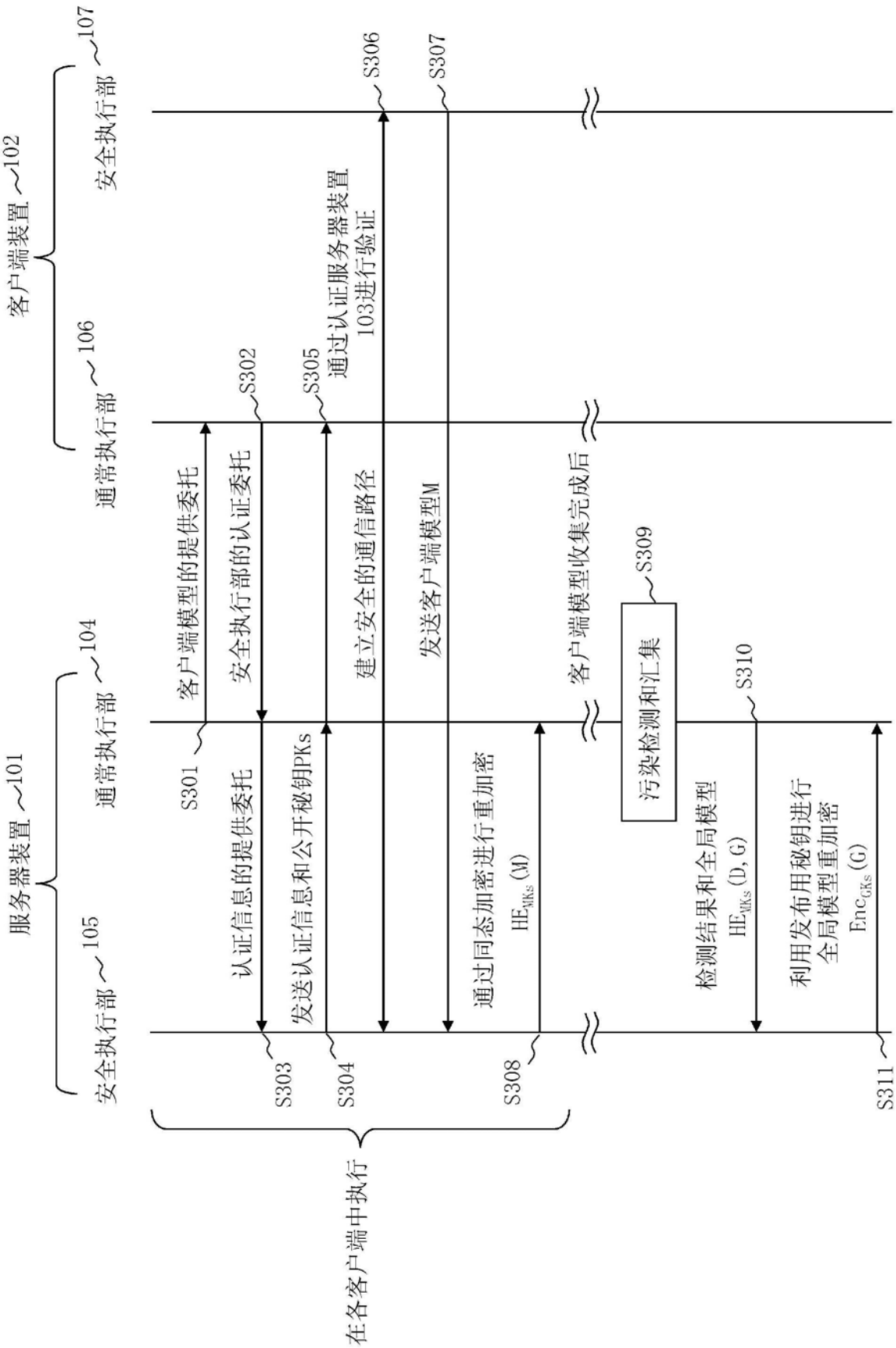

客户端装置 102
服务器装置 101
107
安全执行部
106
通常执行部
105
安全执行部
104
通常执行部
客户端模型的提供委托
S301
S302
安全执行部的认证委托
S305
通过认证服务器装置
103进行验证
S306
S307
建立安全的通信路径
发送客户端模型M
认证信息的提供委托
发送认证信息和公开秘钥PKs
通过同态加密进行重加密
$HE_{MKs}(M)$
S303
S304
S308
客户端模型收集完成后
S309
污染检测和汇集
S310
检测结果和全局模型
$HE_{MKs}(D, G)$
利用发布用秘钥进行
全局模型重加密
$Enc_{GKs}(G)$
S311
在各客户端中执行

图10

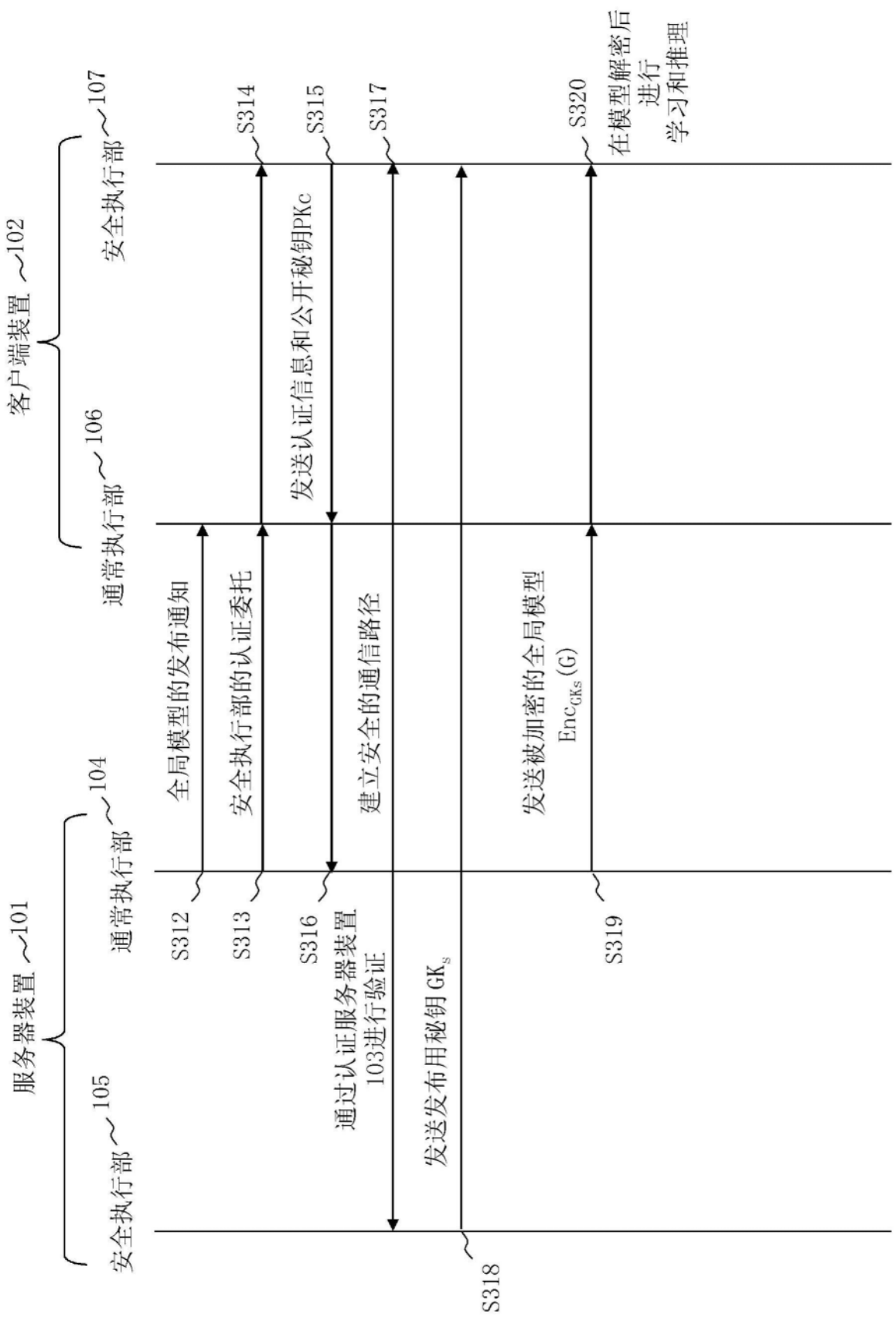

服务器装置 101
安全执行部 105
通常执行部 104
客户端装置 102
通常执行部 106
安全执行部 107
全局模型的发布通知
S312
安全执行部的认证委托
S313
通过认证服务器装置103进行验证
S316
发送发布用秘钥 GK_S
S318
发送认证信息和公开秘钥PKc
S314
S315
建立安全的通信路径
S317
发送被加密的全局模型 $Enc_{GKS}(G)$
S319
S320
在模型解密后进行学习和推理

图11

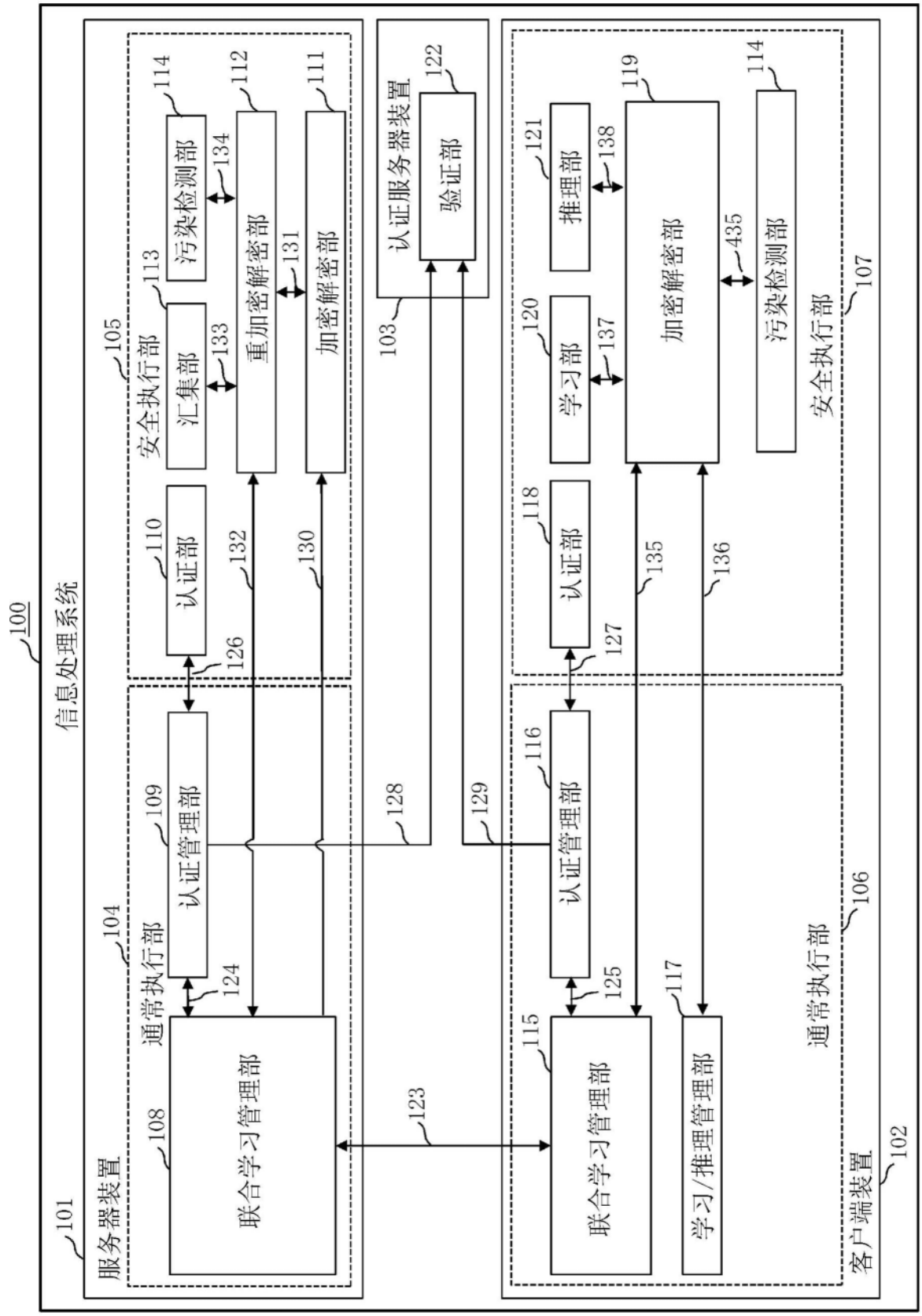

100
信息处理系统
101
服务器装置
104
通常执行部
108
联合学习管理部
109
认证管理部
124
105
安全执行部
110
认证部
126
113
汇集部
133
114
污染检测部
134
112
重加密解密部
131
111
加密解密部
132
130
103
认证服务器装置
122
验证部
128
129
123
102
客户端装置
106
通常执行部
115
联合学习管理部
116
认证管理部
125
117
学习/推理管理部
107
安全执行部
118
认证部
127
120
学习部
137
121
推理部
138
119
加密解密部
135
136
435
114
污染检测部

图12