

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织  
国际局

(43) 国际公布日  
2024 年 7 月 18 日 (18.07.2024)



(10) 国际公布号  
**WO 2024/149202 A1**

- (51) 国际专利分类号:  
**G06F 21/14** (2013.01)
- (21) 国际申请号: PCT/CN2024/071178
- (22) 国际申请日: 2024 年 1 月 8 日 (08.01.2024)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:  
202310077523.5 2023 年 1 月 13 日 (13.01.2023) CN
- (71) 申请人: 维沃移动通信有限公司 (**VIVO MOBILE COMMUNICATION CO., LTD.**) [CN/CN]; 中国广东省东莞市长安镇维沃路 1 号, Guangdong 523863 (CN)。
- (72) 发明人: 洪文博 (**HONG, Wenbo**); 中国广东省东莞市长安镇维沃路 1 号, Guangdong 523863 (CN)。

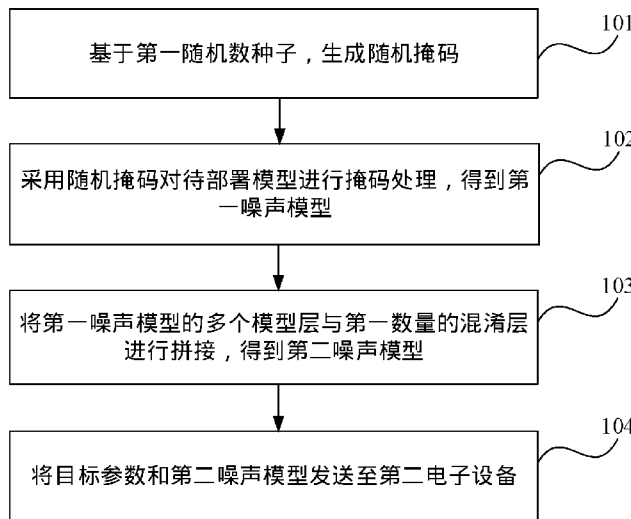
苏涛 (**SU, Tao**); 中国广东省东莞市长安镇维沃路 1 号, Guangdong 523863 (CN)。胡志远 (**HU, Zhiyuan**); 中国广东省东莞市长安镇维沃路 1 号, Guangdong 523863 (CN)。

(74) 代理人: 北京润泽恒知识产权代理有限公司 (**BEIJING RUN ZEHENG INTELLECTUAL PROPERTY LAW FIRM**); 中国北京市海淀区中关村南大街甲 18 号北京国际 C 座 6 层 606, Beijing 100081 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA,

(54) **Title:** MODEL DEPLOYMENT METHOD AND APPARATUS, AND ELECTRONIC DEVICE

(54) 发明名称: 模型部署的方法、装置及电子设备



- 101 Generate a random mask on the basis of a first random seed
- 102 Perform, using the random mask, mask processing on a model to be deployed, to obtain a first noise model
- 103 Stitch a plurality of model layers of the first noise model with a first number of confusion layers to obtain a second noise model
- 104 Send a target parameter and the second noise model to a second electronic device

图 1

(57) **Abstract:** A model deployment method and apparatus, and an electronic device, relating to the technical field of artificial intelligence. The method is applied to a first electronic device, and comprises: generating a random mask on the basis of a first random seed (101); performing, using the random mask, mask processing on a model to be deployed, to obtain a first noise model (102); stitching a plurality of model layers of the first noise model with a first number of confusion layers to obtain a second noise model (103), wherein the number of confusion layers in the second noise model is equal to an element in a first set, and the first set comprises a first number of random numbers generated on the basis of a second random seed; and sending a target parameter and the second noise model to a second electronic device (104), wherein the second electronic device removes the first number of confusion layers in the second noise model on the basis of the target parameter so as to obtain a first noise model (202), and removes the random mask in the first noise model to obtain a model to be deployed (203).

WO 2024/149202 A1

PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:  
— 包括国际检索报告(条约第21条(3))。

(57) 摘要: 一种模型部署的方法、装置及电子设备, 属于人工智能技术领域。方法应用于第一电子设备, 包括: 基于第一随机数种子, 生成随机掩码 (101); 采用随机掩码对待部署模型进行掩码处理, 得到第一噪声模型 (102); 将第一噪声模型的多个模型层与第一数量的混淆层进行拼接, 得到第二噪声模型 (103), 每一混淆层在第二噪声模型中的层数与第一集合中的一个元素相等, 第一集合包括基于第二随机数种子生成的第一数量的随机数; 将目标参数和第二噪声模型发送至第二电子设备 (104), 其中, 第二电子设备基于目标参数去除第二噪声模型中的第一数量的混淆层, 从而得到第一噪声模型 (202), 并去除第一噪声模型中的随机掩码, 得到待部署模型 (203)。

## 模型部署的方法、装置及电子设备

相关申请的交叉引用

本申请要求在 2023 年 1 月 13 日提交中国专利局、申请号为 202310077523.5、名称为“模型部署的方法、装置及电子设备”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

### 技术领域

本申请属于人工智能技术领域，具体涉及一种模型部署的方法、装置及电子设备。

### 背景技术

随着人工智能技术以及模型压缩等技术的快速发展和普及，越来越多的人工智能算法模型集成在了应用中，以进行端侧的算法推断或者训练，这就使得人工智能算法模型具有很大的被窃取风险。

目前，为降低模型文件被窃取的风险，大多采用加密手段。即采用加密算法对模型文件进行加密，模型文件在真正被使用之前处于加密状态。所以，加密状态的模型文件即使被窃取，对方也无法直接使用。

然而，上述方式安全性一般，同时模型文件的应用过程中需要进行加密和解密，造成的延时和开销较大。

### 发明内容

本申请实施例的目的是提供一种模型部署的方法、装置及电子设备，能够解决相关技术中采用加解密方式进行模型保护造成的延时和开销较大的问题。

第一方面，本申请实施例提供了一种模型部署的方法，应用于第一电子设备，该方法包括：

基于第一随机数种子，生成随机掩码；

采用所述随机掩码对待部署模型进行掩码处理，得到第一噪声模型；

将所述第一噪声模型的多个模型层与第一数量的混淆层进行拼接，得到第二噪声模型，其中，所述混淆层为预设模型库中的模型层，每一所述混淆层在所述第二噪声模型中的层数与第一集合中的一个元素相等，所述第一集合包括基于第二随机数种子生成的第一数量的随机数；

将目标参数和所述第二噪声模型发送至第二电子设备，其中，所述第二电子设备基于所述目标参数去除所述第二噪声模型中的所述第一数量的混淆层，从而得到所述第一噪声模型，并去除所述第一噪声模型中的所述随机掩码，得到所述待部署模型；

其中，所述目标参数包括所述第一随机数种子和所述第二随机数种子。

第二方面，本申请实施例提供了一种模型部署的方法，应用于第二电子设备，该方法包括：

接收第一电子设备发送的目标参数和第二噪声模型，其中，所述目标参数包括第一随机数种子和第二随机数种子；所述第二噪声模型为第一噪声模型的多个模型层与第一数量的混淆层进行拼接得到的模型；所述第一噪声模型为采用随机掩码对待部署模型进行掩码处理得到的模型，所述随机掩码为基于所述第一随机数种子生成的掩码；所述混淆层为预设模型库中的模型层，每一所述混淆层在所述第二噪声模型中的层数与第一集合中的一个元素相等，所述第一集合包括基于所述第二随机数种子生成的第一数量的随机数；

基于所述目标参数去除所述第二噪声模型中的所述第一数量的混淆层，得到所述第一噪声模型；

基于所述目标参数去除所述第一噪声模型中的所述随机掩码，得到所述待部署模型。

第三方面，本申请实施例提供了一种模型部署的装置，应用于第一电子设备，该装置包括：

随机数种子模块，用于基于第一随机数种子，生成随机掩码；

掩码模块，用于采用所述随机掩码对待部署模型进行掩码处理，得到第一噪声模型；

拼接模块，用于将所述第一噪声模型的多个模型层与第一数量的混淆层进行拼接，得到第二噪声模型，其中，所述混淆层为预设模型库中的模型层，每一所述混淆层在所述第二噪声模型中的层数与第一集合中的一个元素相等，所述第一集合包括基于第二随机数种子生成的第一数量的随机数；

发送模块，用于将目标参数和所述第二噪声模型发送至第二电子设备，其中，所述第二电子设备基于所述目标参数去除所述第二噪声模型中的所述第一数量的混淆层，从而得到所述第一噪声模型，并去除所述第一噪声模型中的所述随机掩码，得到所述待部署模型；

其中，所述目标参数包括所述第一随机数种子和所述第二随机数种子。

第四方面，本申请实施例提供了一种模型部署的装置，应用于第二电子设备，该装置包括：

接收模块，用于接收第一电子设备发送的目标参数和第二噪声模型，其中，所述目标参数包括第一随机数种子和第二随机数种子；所述第二噪声模型为第一噪声模型的多个模型层与第一数量的混淆层进行拼接得到的模型；所述第一噪声模型为采用随机掩码对待部署模型进行掩码处理得到的模型，所述随机掩码为基于所述第一随机数种子生成的掩码；所述混淆层为预设模型库中的模型层，每一所述混淆层在所述第二噪声模型中的层数与第一集合中的一个元素相等，所述第一集合包括基于所述第二随机数种子生成的第一数量的随机数；

第一去噪模块，用于基于所述目标参数去除所述第二噪声模型中的所述第一数量的混淆层，得到所述第一噪声模型；

第二去噪模块，用于基于所述目标参数去除所述第一噪声模型中的所述随机掩码，得到所述待部署模型。

第五方面，本申请实施例提供了一种电子设备，该电子设备包括处理器和存储器，所述存储器存储可在所述处理器上运行的程序或指令，所述程序或指令被所述处理器执行时实现如第一方面或第二方面所述的方法的步骤。

第六方面，本申请实施例提供了一种可读存储介质，所述可读存储介质上存储程序或指令，所述程序或指令被处理器执行时实现如第一方面或第二方面所述的方法的步骤。

第七方面，本申请实施例提供了一种芯片，所述芯片包括处理器和通信接口，所述通信接口和所述处理器耦合，所述处理器用于运行程序或指令，实现如第一方面或第二方面所述的方法。

第八方面，本申请实施例提供一种计算机程序产品，该程序产品被存储在存储介质中，该程序产品被至少一个处理器执行以实现如第一方面或第二方面所述的方法。

第九方面，本申请实施例提供一种电子设备，所述电子设备被配置成用于执行如第一方面或第二方面所述的方法。

本申请实施例中，模型部署的方法应用于作为模型提供方的第一电子设备。第一电子设备在提供模型之前，利用第一随机数种子生成随机掩码，进而采用随机掩码对待部署模型进行掩码处理，通过往模型中添加噪声以破坏模型，实现对模型的第一层保护。然后将第一噪声模型的多个模型层与第一数量的混淆层进行拼接，通过冗余的混淆层以破坏模型结构，实现对模型的第二层保护。本申请实施例不仅通过两层保护提升了模型的安全性，同时解决了模型加解密造成的延时和开销较大的问题。

#### 附图说明

图1是本申请实施例提供的模型部署的方法的步骤流程图之一；

图2是本申请实施例提供的模型部署的方法的步骤流程图之二；

图3是本申请实施例提供的模型部署的方法的实际应用架构图；

图4是本申请实施例提供的模型部署的装置的结构框图之一；

图5是本申请实施例提供的模型部署的装置的结构框图之二；

图6是本申请实施例提供的电子设备的硬件结构示意图之一；

图7是本申请实施例提供的电子设备的硬件结构示意图之二。

#### 具体实施例

下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清楚地描述，显然，所描述的实施例是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员获得的所有其他实施例，都属于本申请保护的范围。

本申请的说明书和权利要求书中的术语“第一”、“第二”等是用于区别类似的对象，而不用来描述特定的顺序或先后次序。应该理解这样使用的数据在适当情况下可以互换，以便本申请的实施例能够以除了在这里图示或描述的那些以外的顺序实施，且“第一”、“第二”等所区分的对象通常为一类，并不限定对象的个数，例如第一对象可以是一个，也可以是多。此外，说明书以及权利要求中“和/或”表示所连接对象的至少其中之一，

下面结合附图，通过具体的实施例及其应用场景对本申请实施例提供的模型部署的方法进行详细地说明。

如图 1 所示，为本申请实施例提供的模型部署的方法，应用于第一电子设备，其中，第一电子设备作为模型的提供方，存储有待部署模型。

该模型部署的方法，可以包括：

步骤 101：基于第一随机数种子，生成随机掩码。

本步骤中，第一随机数种子为预先设置的一随机数种子，利用该随机数种子按照一预设的伪随机生成函数可以生成多个随机数，随机掩码为该多个随机数组合而成的掩码，其可以为张量或者比特流。可以理解的是，在随机数种子和伪随机生成函数固定的情况下，生成的随机数以及随机数组成的掩码也将是固定的。因此，上述生成的随机数也可以称之为伪随机数，生成的掩码称之为随机掩码。

步骤 102：采用随机掩码对待部署模型进行掩码处理，得到第一噪声模型。

应当说明的是，掩码是深度学习中的常见操作。简单而言，其相当于在原始数据上盖上一层掩膜，从而屏蔽或选择一些特定元素，但不限于此。还可以是对原始数据添加噪声，也就是说，掩码处理即为掩码加噪。

在对待部署模型进行掩码处理时，可以针对模型整体进行掩码处理。例如先将模型本身转换为二进制比特流，然后对二进制比特流就行掩码处理。当然，也可以针对模型的部分数据进行掩码处理。例如从模型中解析出模型权重，进而对模型权重进行掩码处理。可以理解的是，在对模型权重进行掩码处理的情况下，原始模型权重采用随机掩码进行掩码加噪之后将生成新的模型权重，该过程为一可逆的过程，即利用新的模型权重和随机掩码可以进行去掩码处理或掩码去噪，从而得到原始模型权重。其中，原始模型权重为待部署模型的模型权重。新的模型权重为第一噪声模型的模型权重。例如，随机掩码可以为一噪声参数，原始模型权重采用随机掩码进行掩码加噪生成新的模型参数，即为采用一固定算法对原始模型参数和噪声参数进行正向计算，得到正向计算结果，该正向计算结果就是新的模型参数。利用新的模型权重和随机掩码可以进行去掩码处理或掩码去噪生成原始模型权重，即为采用一该固定算法对计算结果和噪声参数进行反向计算，得到反向计算结果，该反向计算结果就是原始模型参数。

值得注意的是，在对待部署模型的模型权重进行掩码处理之后，该待部署模型的模型权重中将包含有随机掩码产生的噪声，且其它模型参数不会发生变化。为便于区分，这里

将掩码处理后的待部署模型称之为第一噪声模型。该第一噪声模型与待部署模型的区别仅在于模型权重不同，第一噪声模型的模型权重为待部署模型的模型权重进行掩码处理后的结果。较佳地，可以采用随机掩码对待部署模型全部或部分的模型权重进行掩码处理，得到第一噪声模型。

5        步骤 103：将第一噪声模型的多个模型层与第一数量的混淆层进行拼接，得到第二噪声模型。

应当说明的是，第二噪声模型中的混淆层用于破坏第一噪声模型或待部署模型的模型结构，对盗取方产生混淆。因此，混淆层并不属于第一噪声模型或待部署模型。这里，混淆层为预设模型库中的模型层，预设模型库为预先设置的存储有多个模型或模型层的数据库。其中，预设模型库中的模型层可以包括待部署模型的模型层和/或预先生成的多个合理的模型层，该合理的模型层为可以与第一噪声模型的模型层实现拼接的模型层。较佳地，混淆层的模型权重可以为随机生成的权重，但不限于此。

可以理解的是，待部署模型包括多个模型层，在拼接混淆层时，可以采用任意方式将混淆层拼接到待部署模型中，使得混淆层作为拼接后模型的模型层。例如可以在待部署模型的每一模型层后面拼接一个混淆层，但不限于此。也可以预先确定各混淆层在拼接后模型中的位置，然后基于确定的位置将各混淆层拼接到待部署模型中。这种情况下，每一混淆层在第二噪声模型中的层数与第一集合中的一个元素相等，第一集合包括基于第二随机数种子生成的第一数量的随机数。混淆层的第一数量可以为一个或者至少两个。这里以第一数量为两个进行说明。例如，基于第二随机数种子生成的两个随机数分别为 3 和 6，则第一集合包括元素 3 和元素 6。相应的，第二噪声模型中的第三层模型层和第六层模型层均为混淆层。

具体的，基于第二随机数种子生成的第一数量的随机数，在预设模型库中随机选择第一数量的模型层作为混淆层。然后按照第一数量的随机数所指示的层数，将混淆层拼接至第一噪声模型中，生成第二噪声模型。可以理解的是，第二随机数种子生成的随机数的取值范围为 $[1, m]$ ，其中， $m$ 等于待部署模型的层数与第一数量的和。

25        步骤 104：将目标参数和第二噪声模型发送至第二电子设备，其中，第二电子设备基于目标参数去除第二噪声模型中的第一数量的混淆层，从而得到第一噪声模型，并去除第一噪声模型中的随机掩码，得到待部署模型。

应当说明的是，目标参数包括第一随机数种子和第二随机数种子。由于通过随机数种子生成的随机掩码和随机数为固定的数据内容，因此，作为模型使用方的第二电子设备可以基于第一随机数种子，生成步骤 101 中的随机掩码；可以基于第二随机数种子，确定混淆层在第二噪声模型中的位置。进而第二电子设备可以去除第二噪声模型中的第一数量的混淆层得到第一噪声模型，并基于生成的随机掩码对第一噪声模型的模型权重进行去掩码或去噪声操作，得到待部署模型

本申请实施例中，模型部署的方法应用于作为模型提供方的第一电子设备。第一电子设备在提供模型之前，利用第一随机数种子生成随机掩码，进而采用随机掩码对待部署模型进行掩码处理，通过往模型中添加噪声以破坏模型，实现对模型的第一层保护。然后将第一噪声模型的多个模型层与第一数量的混淆层进行拼接，通过冗余的混淆层以破坏模型结构，实现对模型的第二层保护。本申请实施例不仅通过两层保护提升了模型的安全性，同时解决了模型加解密造成的延时和开销较大的问题。

可选地，采用随机掩码对待部署模型进行掩码处理，得到第一噪声模型，包括：

采用随机掩码对待部署模型的全部模型权重进行掩码处理，得到第一噪声模型。

应当说明的是，通常情况下神经网络模型的层数较多，且每层可能包括多个神经元。

而每个神经元具有一个模型权重，因此神经网络模型通常具有较多的模型权重，其数量可能是几十上百个，但不限于此。可以理解的是，进行掩码处理后的模型权重，可以视为含有噪声的模型权重。在需要使用待部署模型时，需要去除含有噪声的模型权重中的噪声，这部分模型参数的数量越多，则需要的操作将越复杂、模型的破解难度越大。相应的，模型的安全性也将越高。因此，为尽可能高的提升模型的安全性，将对待部署模型全部的模型权重进行掩码处理。

本申请实施例中，通过对待部署模型的全部模型权重进行掩码处理，可以增加模型的破解难度，进而提升模型的安全性。

可选地，目标参数还包括：由第一字符和第二字符构建的筛选掩码，其中，筛选掩码的每一字符对应待部署模型的一个模型权重。

采用随机掩码对待部署模型进行掩码处理，得到第一噪声模型，包括：

采用随机掩码对待部署模型的目标模型权重进行掩码处理，得到第一噪声模型；

其中，目标模型权重为待部署模型中的对应第一字符的模型权重。

应当说明的是，在采用随机掩码对待部署模型的部分模型权重进行掩码处理的情况下，会仅对一部分模型参数进行掩码处理，而另一部分模型参数不会被进行掩码处理。

这里，第一字符用于表征需要掩码处理，即需要对第一字符对应的模型权重进行掩码处理。第二字符用于表征无需掩码处理，即无需对第二字符对应的模型权重进行掩码处理。示例性地，筛选掩码为多个0和1构成的掩码，0和1的总数量等于模型权重的总数量，每一个0或1均对应待部署模型的一个模型权重。

例如，第一字符可以为1，第二字符可以为0，则筛选掩码01用于表示需要对1对应的模型权重进行掩码处理，而无需对0对应的模型权重进行掩码处理。

可以理解的，筛选掩码为用户基于喜好或经验预先存储至第一电子设备的掩码。也就是说，用户可以自由选择需要进行掩码处理的模型权重，从而增加含有噪声的模型参数的不确定性。当然，还可以采用L1范数（L1-Norm）筛选一些重要的模型参数，基于重要的模型参数构造筛选掩码。具体地，将筛选掩码中对应重要的模型参数的字符设置为第一字符，剩余位设置为第二字符。

本申请实施例中，选择待部署模型中的部分模型参数进行掩码处理，增加了含有噪声的模型参数的不确定性，进而增加了模型的破解难度，提升了模型的安全性。

可选地，目标参数还包括：第三随机数种子，采用随机掩码对待部署模型进行掩码处理，得到第一噪声模型，包括：

5 采用随机掩码对第二数量的目标模型层的模型权重进行掩码处理，得到第一噪声模型。

其中，每一目标模型层在待部署模型中的层数与第二集合中的一个元素相等，第二集合包括基于第三随机数种子生成的第二数量的随机数组成。

应当说明的是，在采用随机掩码对待部署模型部分的模型权重进行掩码处理的情况下，将选择一部分模型参数进行掩码处理，另一部分不进行掩码处理。这里以模型层为维度，  
10 选择一部分模型层的模型参数进行掩码处理，另一部分模型层的模型参数不进行掩码处理。具体的，基于第三随机数种子生成的第二数量的随机数，将第二数量的随机数所指示的模型层确定为目标模型层，进而采用随机掩码对目标模型层的模型权重进行掩码处理，得到第一噪声模型。可以理解的是，第三随机数种子生成的随机数的取值范围为 $[1, n]$ ，其中， $n$ 等于待部署模型的层数。第二数量小于待部署模型的层数。较佳地，第二数量还大于等于  
15 待部署模型的层数的一半。

本申请实施例中，选择待部署模型中的部分模型参数进行掩码处理，增加了含有噪声的模型参数的不确定性，进而增加了模型的破解难度，提升了模型的安全性。

可选的，目标参数还包括：第二噪声模型的数字签名。

应当说明的是，数字签名是只有信息的发送者才能产生的别人无法伪造的一段数字串，  
20 这段数字串同时也是对信息的发送者发送信息真实性的一个有效证明。第二电子设备可以对数字签名进行验签，以确定自身接收到的第二噪声模型是否被篡改。关于数字签名和验签，这里不再详述。

本申请实施例中，通过在目标参数中添加第二噪声模型的数字签名，方便第二电子设备确定第二噪声模型在传输过程中是否被篡改。

25 可选地，将目标参数和第二噪声模型发送至第二电子设备，包括：

将第二噪声模型发送至第二电子设备；

在接收到第二电子设备发送的目标请求的情况下，将目标参数发送至第二电子设备；

其中，目标请求为第二电子设备在本地处于可信执行环境下发送的用于请求目标密钥的请求，目标密钥包括所述目标参数。

30 应当说明的是，目标参数可以理解为待部署模型的密钥，第二噪声模型可以理解为对待部署模型使用密钥添加噪声后的模型。因此，可以将密钥与第二噪声模型一同发送至第二电子设备，以使第二电子设备可以基于接收到数据还原出待部署模型。当然为提升安全性，还可以将两者分批发送，且优先发送第二噪声模型。然后基于第二电子设备返回的目标请求发送目标参数。较佳地，在发送目标参数时，可以先对目标参数进行加密处理，以  
35 密文的方式传输目标参数。例如采用密钥交换协议发送目标参数，但不限于此。

本申请实施例中，将目标参数和第二噪声模型分批发送，在发送第二噪声模型之后，基于第二电子设备的目标请求，向第二电子设备发送目标参数，使得模型部署的方式更加灵活。

如图 2 所示，为本申请实施例提供的模型部署的方法，应用于第二电子设备，其中，第二电子设备作为模型的接收方，用于使用待部署模型进行模型训练或线上使用。

该模型部署的方法，可以包括：

步骤 201：接收第一电子设备发送的目标参数和第二噪声模型。

应当说明的是，目标参数包括第一随机数种子和第二随机数种子；第二噪声模型为第一噪声模型的多个模型层与第一数量的混淆层进行拼接得到的模型；第一噪声模型为采用随机掩码对待部署模型进行掩码处理得到的模型，随机掩码为基于第一随机数种子生成的掩码；混淆层为预设模型库中的模型层，每一混淆层在第二噪声模型中的层数与第一集合中的一个元素相等，第一集合包括基于第二随机数种子生成的第一数量的随机数。可以理解的是，第一电子设备生成并发送目标参数和第二噪声模型的过程、目标参数、第一随机数种子、第二随机数种子、混淆层、第一噪声模型以及随机掩码的相关内容，可以参见图 1 所示实施例，在此不再赘述。

步骤 202：基于目标参数去除第二噪声模型中的第一数量的混淆层，得到第一噪声模型。

应当说明的是，第一数量的混淆层为第二噪声模型中的冗余结构。在第一电子设备侧，将第一噪声模型的模型层与第一数量的混淆层进行拼接，得到第二噪声模型。因此，在第二电子设备侧，将以逆向的处理方式将第二噪声模型还原为第一噪声模型。

值得注意的是，无论在第一电子设备侧还是在第二电子设备侧，基于第二随机数种子生成的随机数相同，即均可以生成第一数量的随机数。例如，在第一电子设备侧，基于第二随机数种子生成的两个随机数分别为 3 和 6。在第二电子设备侧，基于第二随机数种子将同样生成 3 和 6 两个随机数。

本步骤中，将基于第二随机数种子生成的第一数量的随机数，然后按照第一数量的随机数去除第二噪声模型相应的模型层，得到第一噪声模型。假设在第一电子设备侧，基于第二随机数种子生成的两个随机数分别为 3 和 6，则第二噪声模型中的第三层模型层和第六层模型层均为混淆层。在本步骤中，在第二电子设备侧，再次基于第二随机数种子将生成两个随机数 3 和 6。去除第二噪声模型的第三层模型层和第六层模型层，可以得到第一噪声模型。

步骤 203：基于目标参数去除第一噪声模型中的随机掩码，得到待部署模型。

应当说明的是，在第一电子设备侧，采用随机掩码对待部署模型的模型权重进行掩码处理，得到第一噪声模型。因此，在第二电子设备侧，将以逆向的处理方式将第一噪声模型还原为待部署模型。

值得注意的是，无论在第一电子设备侧还是在第二电子设备侧，基于第一随机数种子生成的随机掩码相同。本步骤中，将基于第一随机数种子生成的随机掩码，然后基于生成

的随机掩码对第一噪声模型中的模型权重去掩码处理或掩码去噪，得到待部署模型。由于随机掩码相同，因此，通过逆向的去掩码处理可以去除第一噪声模型中的模型权重中的噪声，进而将第一噪声模型还原为待部署模型。

本申请实施例中，利用第一随机数种子生成随机掩码，进而采用随机掩码对待部署模型进行掩码处理，通过往模型中添加噪声以破坏模型，实现对模型的第一层保护。然后将第一噪声模型的多个模型层与第一数量的混淆层进行拼接，通过冗余的混淆层以破坏模型结构，实现对模型的第二层保护。最后利用目标参数对第二噪声模型进行逆向处理，将其恢复为待部署模型以供使用。本申请实施例不仅通过两层保护提升了模型的安全性，同时解决了模型加解密造成的延时和开销较大的问题。

可选地，目标参数还包括：由第一字符和第二字符构建的筛选掩码，其中，筛选掩码的每一字符对应待部署模型的一个模型权重。

基于目标参数去除第一噪声模型中的随机掩码，得到待部署模型，包括：

基于筛选掩码，确定第一噪声模型中的第一噪声模型权重，其中，第一噪声模型权重在第一噪声模型中的位置与目标模型权重在待部署模型中的位置相同，目标模型权重为待部署模型中的对应第一字符的模型权重。

基于第一随机数种子，生成随机掩码。

基于随机掩码对第一噪声模型中的第一噪声模型权重进行去掩码处理，得到待部署模型。

应当说明的是，在第一电子设备侧基于随机掩码和筛选掩码对待部署模型进行掩码的过程，这里不再赘述。本步骤中，若检测到目标参数中存在筛选掩码，将基于筛选掩码，确定含有噪声的模型权重，即第一噪声模型权重。进而基于目标参数中的第一随机数种子生成的随机掩码，对第一噪声模型权重进行去掩码处理，将第一噪声模型还原为待部署模型。

本申请实施例中，可以选择待部署模型中的部分模型参数进行掩码处理，利用筛选掩码对部分模型权重添加噪声，同时利用筛选掩码确定含有噪声的模型权重，可以实现模型的破坏与还原，提升了模型的安全性。

可选地，目标参数还包括：第三随机数种子。

基于目标参数去除第一噪声模型中的随机掩码，得到待部署模型，包括：

基于第三随机数种子，生成第二数量的随机数。

基于第二数量的随机数，确定第一噪声模型中第二数量的噪声模型层；其中，每一噪声模型层在第一噪声模型中的层数与第二数量的随机数中的一个随机数相等。

基于第一随机数种子，生成随机掩码。

基于随机掩码对第二噪声模型权重进行去掩码处理，得到待部署模型；其中，第二噪声模型权重为第二数量的噪声模型层的模型权重。

应当说明的是，在第一电子设备侧基于随机掩码和混淆层对待部署模型进行处理的过程，这里不再赘述。本步骤中，若检测到目标参数中含有第三随机数种子，将基于第三随机数种子生成第二数量的随机数，然后将第二数量的随机数指示的模型层确定为含有噪声或进行过掩码处理的模型层，即噪声模型层。进而基于目标参数中的第一随机数种子生成的随机掩码，对噪声模型层的模型权重进行去掩码处理，将第一噪声模型还原为待部署模型。

值得注意的是，无论在第一电子设备侧还是在第二电子设备侧，基于第三随机数种子生成的随机数相同，即均可以生成第二数量的随机数。例如，在第一电子设备侧，基于第三随机数种子生成的两个随机数分别为3和6。在第二电子设备侧，基于第三随机数种子将同样生成3和6两个随机数。

本申请实施例中，可以选择待部署模型中的部分模型参数进行掩码处理，利用第三随机数种子对部分模型权重添加噪声，同时利用第三随机数种子确定含有噪声的模型权重，可以实现模型的破坏与还原，提升了模型的安全性。

可选地，目标参数还包括：第二噪声模型的数字签名；

基于目标参数去除第二噪声模型中的第一数量的混淆层，得到第一噪声模型，包括：在数字签名通过验签的情况下，基于目标参数去除第二噪声模型中的第一数量的混淆层，得到第一噪声模型。

应当说明的是，数字签名是只有信息的发送者才能产生的别人无法伪造的一段数字串，这段数字串同时也是对信息的发送者发送信息真实性的一个有效证明。第二电子设备可以对数字签名进行验签，以确定自身接收到的第二噪声模型是否被篡改。关于数字签名和验签，这里不再详述。

本申请实施例中，通过在目标参数中添加第二噪声模型的数字签名，方便第二电子设备确定第二噪声模型在传输过程中是否被篡改。

可选地，接收第一电子设备发送的目标参数和第二噪声模型，包括：

接收第一电子设备发送的第二噪声模型；

在第二电子设备本地处于可信执行环境下，发送用于请求目标密钥的第一请求，其中，目标密钥包括目标参数；

接收第一电子设备基于第一请求返回的目标密钥。

应当说明的是，网络中的威胁包括：半诚实威胁和恶意威胁。其中，在半诚实威胁下，攻击者无法攻破设备，只能在预定的规则和有限权限内进行攻击实现自己的目的。在恶意威胁下，攻击者能够完全掌握被攻击的设备，并且不遵守预定的协议，控制整个运行流程。因此，本申请可以基于两种威胁设置两种不同的场景进行应对。在半诚实威胁场景下，由于攻击者是半诚实方，即不会攻破第二电子设备的系统进行干预。可以在应用或服务首次加载时采用简单的加密通信方式将目标参数加载至第二电子设备的私有空间。在恶意威胁场景中，第二电子设备已经被恶意攻击者攻破，攻击者将会获得设备控制权，掌握运行时

信息。此时，需要采用加强方案对目标参数进行分发、管理和使用。具体来说，在应用或服务首次加载时，于可信执行环境中使用密钥交换协议与第一电子设备交互获取目标参数，即目标参数只在可信执行环境中可见。其中，可信执行环境：完全可信的运行环境，无法被攻破和运行时监测，包括现有的 TEE（可信执行环境，Trusted Execution Environment）技术以及机密计算等技术之上搭建的运行环境。

如图 3 所示，本申请实施例还提供了一种模型部署的方法，应用于第一设备和第二设备组成的网络系统，其中，第一设备相当于图 1 所示实施例中的第一电子设备，第二设备相当于图 2 所示实施例中的第二电子设备，在此不再赘述。

该方法可以包括：

步骤 301：在第一设备中生成掩码、冗余结构以及相应密钥。具体的，使用一个预定的随机数种子 key1 生成随机掩码 Mask1。使用预定的随机数种子 key2，生成 k 个目标范围内的随机数，其中目标范围为  $[1, k+n]$ ，其中，n 为待部署模型的层数。在预设模型库中随机选择 k 个模型层作为冗余结构。较佳地，k 小于等于 n，且 k 大于等于目标值，目标值为  $n/4$  和 1 中的最大值。这里，随机掩码 Mask1 的密钥为随机数种子 key1；冗余结构的密钥为随机数种子 key2。

步骤 302：第一设备组合保护模型后下发部署。具体的，使用随机掩码 Mask1 对待部署模型中的模型权重进行掩码处理，然后拼接冗余结构，得到  $k+n$  层的保护模型，保护模型相当于上述实施例中的第二噪声模型。然后将保护模型和保护模型的数字签名作为打包文件下发至第二设备。

步骤 303：第一设备进行密钥下发。具体的，在接收到第二设备发送的目标请求的情况下，将随机数种子 key1 和随机数种子 key2 作为密钥发送至第二设备。

步骤 304：第二设备在接收到打包文件后，对保护模型进行文件部署。

步骤 305：请求第一设备下发密钥，将密钥存储至可信执行空间中。该可信执行空间为第二设备中具有可信执行环境的存储空间。

步骤 306：在需要使用模型的情况下，利用打包文件中的数字签名对保护模型进行验签以验证保护模型的完整性。若验签不通过，则请求第一设备重新下发打包文件。若验签通过，则执行步骤 307。

步骤 307：基于可信执行空间中的密钥，去除保护模型中的冗余结构，并进行去掩码处理，得到待部署模型。然后使用待部署模型进行模型训练或数据推断。值得注意的是，在恶意威胁场景中，本步骤将在第二设备的可信执行环境中进行。

应当说明的是，使用随机掩码 Mask1 对待部署模型中的模型权重进行掩码处理时，可以是覆盖整个模型也可以仅覆盖部分模型，即可以对待部署模型全部的模型权重进行掩码处理，也可以对待部署模型的部分模型权重进行掩码处理，根据安全性和性能的需要可以具体采用以下方式。

方式 1：对整个模型加上一层掩码噪声，即对待部署模型全部的模型权重进行掩码处理。

方式 2：通过 L1 范式确定重要的模型权重，随后针对这部分权重（可以大于总权重数的 20%）构造由 0 和 1 组成的筛选掩码 Mask2，最后对 1 对应的模型权重使用随机掩码 Mask1 进行掩码加噪。

方式 3：使用随机数种子 key3 生成又一数值范围内的 n1 个随机数，其中，该数值范围  
5 为  $[1, n]$ ，n 等于待部署模型的层数。n1 小于等于 n，且 n1 大于等于  $n/2$ 。对 n1 个随机数指示的模型层中的模型权重使用随机掩码 Mask1 进行掩码加噪。

因此，第二设备接收到的密钥中通常包含有：随机数种子 key1 和随机数种子 key2。若其还包含有随机数种子 key3，说明第一设备采用了方式 3 进行掩码加噪。在进行掩码去噪时，先基于随机数种子 key3，确定需要去噪的模型层，进而对这部分需要去噪的模型层利  
10 用随机掩码 Mask1 进行掩码去噪。若其还包含有筛选掩码 Mask2，说明第一设备采用了方式 2 进行掩码加噪。在进行掩码去噪时，先基于筛选掩码 Mask2，确定需要去噪的模型权重，进而对这部分需要去噪的模型权重利用随机掩码 Mask1 进行掩码去噪。

本申请实施例中，使用随机掩码技术以较低开销和时延破坏模型完整性和可用性；在进行模型权重保护的同时，使用冗余模型结构进行模型结构混淆保护，且能够迷惑攻击者  
15 使其错误判断所获模型为可用的真实模型。

需要说明的是，本申请实施例提供的模型部署的方法，执行主体可以为模型部署的装置，或者该模型部署的装置中的用于执行模型部署的方法的控制模块。本申请实施例中以模型部署的装置执行模型部署的方法为例，说明本申请实施例提供的模型部署的装置。

如图 4 所示，本申请实施例还提供了一种模型部署的装置，应用于第一电子设备，该  
20 装置包括：

随机数种子模块 41，用于基于第一随机数种子，生成随机掩码。

掩码模块 42，用于采用随机掩码对待部署模型进行掩码处理，得到第一噪声模型。

拼接模块 43，用于将第一噪声模型的多个模型层与第一数量的混淆层进行拼接，得到第二噪声模型，其中，混淆层为预设模型库中的模型层，每一混淆层在第二噪声模型中的  
25 层数与第一集合中的一个元素相等，第一集合包括基于第二随机数种子生成的第一数量的随机数。

发送模块 44，用于将目标参数和第二噪声模型发送至第二电子设备，其中，第二电子设备基于目标参数去除第二噪声模型中的第一数量的混淆层，从而得到第一噪声模型，并去除第一噪声模型中的随机掩码，得到待部署模型。

其中，目标参数包括第一随机数种子和第二随机数种子。

可选地，掩码模块 42，具体用于采用随机掩码对待部署模型的全部模型权重进行掩码处理，得到第一噪声模型。

可选地，目标参数还包括：由第一字符和第二字符构建的筛选掩码，其中，筛选掩码的每一字符对应待部署模型的一个模型权重。

掩码模块 42，具体用于采用随机掩码对待部署模型的目标模型权重进行掩码处理，得到第一噪声模型。

其中，目标模型权重为待部署模型中的对应第一字符的模型权重。

5 可选地，目标参数还包括：第三随机数种子，掩码模块 42，具体用于采用随机掩码对待部署模型的第二数量的目标模型层的模型权重进行掩码处理，得到第一噪声模型。

其中，每一目标模型层在待部署模型中的层数与第二集合中的一个元素相等，第二集合包括基于第三随机数种子生成的第二数量的随机数。

可选地，目标参数还包括：第二噪声模型的数字签名。

可选地，发送模块 44，包括：

10 第一发送单元，用于将第二噪声模型发送至第二电子设备。

第二发送单元，用于在接收到第二电子设备发送的目标请求的情况下，将目标参数发送至第二电子设备。

其中，目标请求为第二电子设备在本地处于可信执行环境下发送的用于请求目标密钥的请求，目标密钥包括所述目标参数。

15 本申请实施例中，第一电子设备在提供模型之前，利用第一随机数种子生成随机掩码，进而采用随机掩码对待部署模型的模型权重进行掩码处理，通过往模型权重中添加噪声以破坏模型权重，实现对模型的第一层保护。然后将第一噪声模型的模型层与第一数量的混淆层进行拼接，通过冗余的混淆层以破坏模型结构，实现对模型的第二层保护。本申请实施例不仅通过两层保护提升了模型的安全性，同时解决了模型加解密造成的延时和开销较大的问题。

20 本申请实施例中的模型部署的装置可以是电子设备，也可以是电子设备中的部件，例如集成电路或芯片。该电子设备可以是终端，也可以为除终端之外的其他设备。示例性的，电子设备可以为手机、平板电脑、笔记本电脑、掌上电脑、车载电子设备、移动上网装置（Mobile Internet Device, MID）、增强现实(augmented reality, AR)/虚拟现实(virtual reality, VR)设备、机器人、可穿戴设备、超级移动个人计算机（ultra-mobile personal computer, UMPC）、上网本或者个人数字助理（personal digital assistant, PDA）等，还可以为服务器、网络附属存储器（Network Attached Storage, NAS）、个人计算机（personal computer, PC）、电视机（television, TV）、柜员机或者自助机等，本申请实施例不作具体限定。

25 本申请实施例中的模型部署的装置可以为具有操作系统的装置。该操作系统可以为安卓（Android）操作系统，可以为 iOS 操作系统，还可以为其他可能的操作系统，本申请实施例不作具体限定。

本申请实施例提供的模型部署的装置能够实现图 1 的方法实施例实现的各个过程，实现相同的技术效果，为避免重复，这里不再赘述。

30 如图 5 所示，本申请实施例还提供了一种模型部署的装置，应用于第二电子设备，该装置包括：

接收模块 51，用于接收第一电子设备发送的目标参数和第二噪声模型，其中，目标参数包括第一随机数种子和第二随机数种子；第二噪声模型为第一噪声模型的多个模型层与第一数量的混淆层进行拼接得到的模型；第一噪声模型为采用随机掩码对待部署模型进行掩码处理得到的模型，随机掩码为基于第一随机数种子生成的掩码；混淆层为预设模型库中的模型层，每一混淆层在第二噪声模型中的层数与第一集合中的一个元素相等，第一集合包括基于第二随机数种子生成的第一数量的随机数。

第一去噪模块 52，用于基于目标参数去除第二噪声模型中的第一数量的混淆层，得到第一噪声模型。

第二去噪模块 53，用于基于目标参数去除第一噪声模型中的随机掩码，得到待部署模型。

可选地，目标参数还包括：由第一字符和第二字符构建的筛选掩码，其中，筛选掩码的每一字符对应待部署模型的一个模型权重。

第二去噪模块 53，包括：

筛选单元，用于基于筛选掩码，确定第一噪声模型中的第一噪声模型权重，其中，第一噪声模型权重在第一噪声模型中的位置与目标模型权重在待部署模型中的位置相同，目标模型权重为待部署模型中的对应第一字符的模型权重。

第一随机数种子单元，用于基于第一随机数种子，生成随机掩码。

第一去噪单元，用于基于随机掩码对第一噪声模型中的第一噪声模型权重进行去掩码处理，得到待部署模型。

可选地，目标参数还包括：第三随机数种子。

第二去噪模块 53，包括：

第二随机数种子单元，用于基于第三随机数种子，生成第二数量的随机数。

确定单元，用于基于第二数量的随机数，确定第一噪声模型中第二数量的噪声模型层；其中，每一噪声模型层在第一噪声模型中的层数与第二数量的随机数中的一个随机数相等。

第三随机数种子单元，用于基于第一随机数种子，生成随机掩码。

第二去噪单元，用于基于随机掩码对第二噪声模型权重进行去掩码处理，得到待部署模型；其中，第二噪声模型权重为第二数量的噪声模型层的模型权重。

可选地，目标参数还包括：第二噪声模型的数字签名。

第一去噪模块 52，具体用于在数字签名通过验签的情况下，基于目标参数去除第二噪声模型中的第一数量的混淆层，得到第一噪声模型。

可选地，接收模块 51，具体用于：

接收第一电子设备发送的所述第二噪声模型；

在所述第二电子设备本地处于可信执行环境下，发送用于请求目标密钥的第一请求，其中，所述目标密钥包括所述目标参数；

接收所述第一电子设备基于所述第一请求返回的所述目标密钥。

本申请实施例中，利用第一随机数种子生成随机掩码，进而采用随机掩码对待部署模型进行掩码处理，通过往模型中添加噪声以破坏模型，实现对模型的第一层保护。然后将第一噪声模型的多个模型层与第一数量的混淆层进行拼接，通过冗余的混淆层以破坏模型结构，实现对模型的第二层保护。最后利用目标参数对第二噪声模型进行逆向处理，将其恢复为待部署模型以供使用。本申请实施例不仅通过两层保护提升了模型的安全性，同时解决了模型加解密造成的延时和开销较大的问题。

本申请实施例中的模型部署的装置可以是电子设备，也可以是电子设备中的部件，例如集成电路或芯片。该电子设备可以是终端，也可以为除终端之外的其他设备。示例性的，电子设备可以为手机、平板电脑、笔记本电脑、掌上电脑、车载电子设备、移动上网装置（Mobile Internet Device, MID）、增强现实(augmented reality, AR)/虚拟现实(virtual reality, VR)设备、机器人、可穿戴设备、超级移动个人计算机（ultra-mobile personal computer, UMPC）、上网本或者个人数字助理（personal digital assistant, PDA）等，还可以为服务器、网络附属存储器（Network Attached Storage, NAS）、个人计算机（personal computer, PC）、电视机（television, TV）、柜员机或者自助机等，本申请实施例不作具体限定。

本申请实施例中的模型部署的装置可以为具有操作系统的装置。该操作系统可以为安卓（Android）操作系统，可以为 iOS 操作系统，还可以为其他可能的操作系统，本申请实施例不作具体限定。

本申请实施例提供的模型部署的装置能够实现图 2 的方法实施例实现的各个过程，实现相同的技术效果，为避免重复，这里不再赘述。

可选地，如图 6 所示，本申请实施例还提供一种电子设备 600，包括处理器 601 和存储器 602，存储器 602 上存储有可在所述处理器 601 上运行的程序或指令，该程序或指令被处理器 601 执行时实现上述模型部署的方法实施例的各个步骤，且能达到相同的技术效果，为避免重复，这里不再赘述。

需要说明的是，本申请实施例中的电子设备包括上述所述的移动电子设备和非移动电子设备。

图 7 为实现本申请实施例的一种电子设备的硬件结构示意图。

该电子设备 700 包括但不限于：射频单元 701、网络模块 702、音频输出单元 703、输入单元 704、传感器 705、显示单元 706、用户输入单元 707、接口单元 708、存储器 709、以及处理器 710 等部件。

本领域技术人员可以理解，电子设备 700 还可以包括给各个部件供电的电源（比如电池），电源可以通过电源管理系统与处理器 710 逻辑相连，从而通过电源管理系统实现管理充电、放电、以及功耗管理等功能。图 7 中示出的电子设备结构并不构成对电子设备的限定，电子设备可以包括比图示更多或更少的部件，或者组合某些部件，或者不同的部件布置，在此不再赘述。

其中，在电子设备为第一电子设备的情况下，处理器 710，用于基于第一随机数种子，生成随机掩码。

处理器 710，还用于采用随机掩码对待部署模型进行掩码处理，得到第一噪声模型。

处理器 710，还用于将第一噪声模型的多个模型层与第一数量的混淆层进行拼接，得到  
5 第二噪声模型，其中，混淆层为预设模型库中的模型层，每一混淆层在第二噪声模型中的  
层数与第一集合中的一个元素相等，第一集合包括基于第二随机数种子生成的第一数量的  
随机数。

网络模块 702，用于将目标参数和第二噪声模型发送至第二电子设备，其中，第二电子  
设备基于目标参数去除第二噪声模型中的第一数量的混淆层，从而得到第一噪声模型，并  
10 去除第一噪声模型中的随机掩码，得到待部署模型。

其中，目标参数包括第一随机数种子和第二随机数种子。

本申请实施例中，第一电子设备在提供模型之前，利用第一随机数种子生成随机掩码，  
进而采用随机掩码对待部署模型进行掩码处理，通过往模型中添加噪声以破坏模型，实现  
对模型的第一层保护。然后将第一噪声模型的多个模型层与第一数量的混淆层进行拼接，  
15 通过冗余的混淆层以破坏模型结构，实现对模型的第二层保护。本申请实施例不仅通过两  
层保护提升了模型的安全性，同时解决了模型加解密造成的延时和开销较大的问题。

在电子设备为第二电子设备的情况下，网络模块 702，用于接收第一电子设备发送的目  
标参数和第二噪声模型，其中，目标参数包括第一随机数种子和第二随机数种子；第二噪  
声模型为第一噪声模型的多个模型层与第一数量的混淆层进行拼接得到的模型；第一噪声  
20 模型为采用随机掩码对待部署模型进行掩码处理得到的模型，随机掩码为基于第一随机数  
种子生成的掩码；混淆层为预设模型库中的模型层，每一混淆层在第二噪声模型中的层数  
与第一集合中的一个元素相等，第一集合包括基于第二随机数种子生成的第一数量的随机  
数。

处理器 710，用于基于目标参数去除第二噪声模型中的第一数量的混淆层，得到第一噪  
25 声模型。

处理器 710，还用于基于目标参数去除第一噪声模型中的随机掩码，得到待部署模型。

本申请实施例中，利用第一随机数种子生成随机掩码，进而采用随机掩码对待部署模  
型进行掩码处理，通过往模型中添加噪声以破坏模型，实现对模型的第一层保护。然后将  
第一噪声模型的多个模型层与第一数量的混淆层进行拼接，通过冗余的混淆层以破坏模型  
30 结构，实现对模型的第二层保护。最后利用目标参数对第二噪声模型进行逆向处理，将其  
恢复为待部署模型以供使用。本申请实施例不仅通过两层保护提升了模型的安全性，同时  
解决了模型加解密造成的延时和开销较大的问题。

应理解的是，本申请实施例中，输入单元 704 可以包括图形处理器（Graphics Processing  
Unit, GPU）7041 和麦克风 7042，图形处理器 7041 对在视频捕获模式或图像捕获模式中由  
35 图像捕获装置（如摄像头）获得的静态图片或视频的图像数据进行处理。显示单元 706 可

包括显示面板 7061, 可以采用液晶显示器、有机发光二极管等形式来配置显示面板 7061。用户输入单元 707 包括触控面板 7071 以及其他输入设备 7072 中的至少一种。触控面板 7071, 也称为触摸屏。触控面板 7071 可包括触摸检测装置和触摸控制器两个部分。其他输入设备 7072 可以包括但不限于物理键盘、功能键(比如音量控制按键、开关按键等)、轨迹球、鼠标、操作杆, 在此不再赘述。

存储器 709 可用于存储软件程序以及各种数据。存储器 709 可主要包括存储程序或指令的第一存储区和存储数据的第二存储区, 其中, 第一存储区可存储操作系统、至少一个功能所需的应用程序或指令(比如声音播放功能、图像播放功能等)等。此外, 存储器 709 可以包括易失性存储器或非易失性存储器, 或者, 存储器 709 可以包括易失性和非易失性存储器两者。其中, 非易失性存储器可以是只读存储器(Read-Only Memory, ROM)、可编程只读存储器(Programmable ROM, PROM)、可擦除可编程只读存储器(Erasable PROM, EPROM)、电可擦除可编程只读存储器(Electrically EPROM, EEPROM)或闪存。易失性存储器可以是随机存取存储器(Random Access Memory, RAM), 静态随机存取存储器(Static RAM, SRAM)、动态随机存取存储器(Dynamic RAM, DRAM)、同步动态随机存取存储器(Synchronous DRAM, SDRAM)、双倍数据速率同步动态随机存取存储器(Double Data Rate SDRAM, DDRSDRAM)、增强型同步动态随机存取存储器(Enhanced SDRAM, ESDRAM)、同步连接动态随机存取存储器(Synch link DRAM, SLDRAM)和直接内存总线随机存取存储器(Direct Rambus RAM, DRRAM)。本申请实施例中的存储器 709 包括但不限于这些和任意其它适合类型的存储器。

处理器 710 可包括一个或多个处理单元; 可选地, 处理器 710 集成应用处理器和调制解调处理器, 其中, 应用处理器主要处理涉及操作系统、用户界面和应用程序等的操作, 调制解调处理器主要处理无线通信信号, 如基带处理器。可以理解的是, 上述调制解调处理器也可以不集成到处理器 710 中。

本申请实施例还提供一种可读存储介质, 所述可读存储介质上存储有程序或指令, 该程序或指令被处理器执行时实现上述模型部署的方法实施例的各个过程, 且能达到相同的技术效果, 为避免重复, 这里不再赘述。

其中, 所述处理器为上述实施例中所述的电子设备中的处理器。所述可读存储介质, 包括计算机可读存储介质, 如计算机只读存储器(Read-Only Memory, ROM)、随机存取存储器(Random Access Memory, RAM)、磁碟或者光盘等。

本申请实施例另提供了一种芯片, 所述芯片包括处理器和通信接口, 所述通信接口和所述处理器耦合, 所述处理器用于运行程序或指令, 实现上述模型部署的方法实施例的各个过程, 且能达到相同的技术效果, 为避免重复, 这里不再赘述。

应理解, 本申请实施例提到的芯片还可以称为系统级芯片、系统芯片、芯片系统或片上系统芯片等。

本申请实施例提供一种计算机程序产品，该程序产品被存储在存储介质中，该程序产品被至少一个处理器执行以实现如上述模型部署的方法实施例的各个过程，且能达到相同的技术效果，为避免重复，这里不再赘述。

需要说明的是，在本文中，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、物品或者装置不仅包括那些要素，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、物品或者装置所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括该要素的过程、方法、物品或者装置中还存在另外的相同要素。此外，需要指出的是，本申请实施方式中的方法和装置的范围不限按示出或讨论的顺序来执行功能，还可包括根据所涉及的功能按基本同时的方式或按相反的顺序来执行功能，例如，可以按不同于所描述的次序来执行所描述的方法，并且还可以添加、省去、或组合各种步骤。另外，参照某些示例所描述的特征可在其他示例中被组合。

通过以上的实施方式的描述，本领域的技术人员可以清楚地了解到上述实施例方法可借助软件加必需的通用硬件平台的方式来实现，当然也可以通过硬件，但很多情况下前者是更佳的实施方式。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分可以以计算机软件产品的形式体现出来，该计算机软件产品存储在一个存储介质（如 ROM/RAM、磁碟、光盘）中，包括若干指令用以使得一台终端（可以是手机，计算机，服务器，或者网络设备等）执行本申请各个实施例所述的方法。

上面结合附图对本申请的实施例进行了描述，但是本申请并不局限于上述的具体实施方式，上述的具体实施方式仅仅是示意性的，而不是限制性的，本领域的普通技术人员在本申请的启示下，在不脱离本申请宗旨和权利要求所保护的范围情况下，还可做出很多形式，均属于本申请的保护之内。

## 权 利 要 求

1. 一种模型部署的方法，应用于第一电子设备，所述方法包括：

基于第一随机数种子，生成随机掩码；

采用所述随机掩码对待部署模型进行掩码处理，得到第一噪声模型；

5 将所述第一噪声模型的多个模型层与第一数量的混淆层进行拼接，得到第二噪声模型，其中，所述混淆层为预设模型库中的模型层，每一所述混淆层在所述第二噪声模型中的层数与第一集合中的一个元素相等，所述第一集合包括基于第二随机数种子生成的第一数量的随机数；

10 将目标参数和所述第二噪声模型发送至第二电子设备，其中，所述第二电子设备基于所述目标参数去除所述第二噪声模型中所述第一数量的混淆层，从而得到所述第一噪声模型，并去除所述第一噪声模型中的所述随机掩码，得到所述待部署模型；

其中，所述目标参数包括所述第一随机数种子和所述第二随机数种子。

2. 根据权利要求1所述的方法，其中，所述采用所述随机掩码对待部署模型进行掩码处理，得到第一噪声模型，包括：

15 采用所述随机掩码对待部署模型的全部模型权重进行掩码处理，得到第一噪声模型。

3. 根据权利要求1所述的方法，其中，所述目标参数还包括：由第一字符和第二字符构建的筛选掩码，其中，所述筛选掩码的每一字符对应所述待部署模型的一个模型权重；

所述采用所述随机掩码对待部署模型进行掩码处理，得到第一噪声模型，包括：

20 采用所述随机掩码对所述待部署模型的目标模型权重进行掩码处理，得到第一噪声模型；

其中，所述目标模型权重为所述待部署模型中的对应所述第一字符的模型权重。

4. 根据权利要求1所述的方法，其中，所述目标参数还包括：第三随机数种子，所述采用所述随机掩码对待部署模型进行掩码处理，得到第一噪声模型，包括：

25 采用所述随机掩码对第二数量的目标模型层的模型权重进行掩码处理，得到第一噪声模型；

其中，每一所述目标模型层在所述待部署模型中的层数与第二集合中的一个元素相等，所述第二集合包括基于所述第三随机数种子生成的第二数量的随机数。

5. 根据权利要求1所述的方法，其中，所述目标参数还包括：所述第二噪声模型的数字签名。

6. 根据权利要求1所述的方法，其中，所述将所述目标参数和所述第二噪声模型发送至第二电子设备，包括：

将所述第二噪声模型发送至所述第二电子设备；

35 在接收到所述第二电子设备发送的目标请求的情况下，将所述目标参数发送至所述第二电子设备；

其中，所述目标请求为所述第二电子设备在本地处于可信执行环境下发送的用于请求目标密钥的请求，所述目标密钥包括所述目标参数。

7. 一种模型部署的方法，应用于第二电子设备，所述方法包括：

接收第一电子设备发送的目标参数和第二噪声模型，其中，所述目标参数包括第一  
5 随机数种子和第二随机数种子；所述第二噪声模型为第一噪声模型的多个模型层与第一数量的混淆层进行拼接得到的模型；所述第一噪声模型为采用随机掩码对待部署模型进行掩码处理得到的模型，所述随机掩码为基于所述第一随机数种子生成的掩码；所述混淆层为预设模型库中的模型层，每一所述混淆层在所述第二噪声模型中的层数与第一集合中的一个元素相等，所述第一集合包括基于所述第二随机数种子生成的第一数量的随  
10 机数；

基于所述目标参数去除所述第二噪声模型中的所述第一数量的混淆层，得到所述第一噪声模型；

基于所述目标参数去除所述第一噪声模型中的所述随机掩码，得到所述待部署模型。

8. 根据权利要求7所述的方法，其中，所述目标参数还包括：由第一字符和第二字  
15 符构建的筛选掩码，其中，所述筛选掩码的每一字符对应所述待部署模型的一个模型权重；

所述基于所述目标参数去除所述第一噪声模型中的所述随机掩码，得到所述待部署模型，包括：

基于所述筛选掩码，确定所述第一噪声模型中的第一噪声模型权重，其中，所述第  
20 一噪声模型权重在所述第一噪声模型中的位置与目标模型权重在所述待部署模型中的位置相同，所述目标模型权重为所述待部署模型中的对应所述第一字符的模型权重；

基于所述第一随机数种子，生成所述随机掩码；

基于所述随机掩码对所述第一噪声模型中的所述第一噪声模型权重进行去掩码处理，得到所述待部署模型。

25 9. 根据权利要求7所述的方法，其中，所述目标参数还包括：第三随机数种子；

所述基于所述目标参数去除所述第一噪声模型中的所述随机掩码，得到所述待部署模型，包括：

基于所述第三随机数种子，生成第二数量的随机数；

基于所述第二数量的随机数，确定所述第一噪声模型中第二数量的噪声模型层；其  
30 中，每一所述噪声模型层在所述第一噪声模型中的层数与所述第二数量的随机数中的一个随机数相等；

基于所述第一随机数种子，生成所述随机掩码；

基于所述随机掩码对第二噪声模型权重进行去掩码处理，得到所述待部署模型；其中，所述第二噪声模型权重为所述第二数量的噪声模型层的模型权重。

10. 根据权利要求 7 所述的方法, 其中, 所述目标参数还包括: 所述第二噪声模型的数字签名;

所述基于所述目标参数去除所述第二噪声模型中的所述第一数量的混淆层, 得到所述第一噪声模型, 包括:

5 在所述数字签名通过验签的情况下, 基于所述目标参数去除所述第二噪声模型中的所述第一数量的混淆层, 得到所述第一噪声模型。

11. 根据权利要求 7 所述的方法, 其中, 接收第一电子设备发送的目标参数和第二噪声模型, 包括:

接收第一电子设备发送的所述第二噪声模型;

10 在所述第二电子设备本地处于可信执行环境下, 发送用于请求目标密钥的第一请求, 其中, 所述目标密钥包括所述目标参数;

接收所述第一电子设备基于所述第一请求返回的所述目标密钥。

12. 一种模型部署的装置, 应用于第一电子设备, 所述装置包括:

随机数种子模块, 用于基于第一随机数种子, 生成随机掩码;

15 掩码模块, 用于采用所述随机掩码对待部署模型进行掩码处理, 得到第一噪声模型;

拼接模块, 用于将所述第一噪声模型的多个模型层与第一数量的混淆层进行拼接, 得到第二噪声模型, 其中, 所述混淆层为预设模型库中的模型层, 每一所述混淆层在所述第二噪声模型中的层数与第一集合中的一个元素相等, 所述第一集合包括基于第二随机数种子生成的第一数量的随机数;

20 发送模块, 用于将目标参数和所述第二噪声模型发送至第二电子设备, 其中, 所述第二电子设备基于所述目标参数去除所述第二噪声模型中的所述第一数量的混淆层, 从而得到所述第一噪声模型, 并去除所述第一噪声模型中的所述随机掩码, 得到所述待部署模型;

其中, 所述目标参数包括所述第一随机数种子和所述第二随机数种子。

25 13. 根据权利要求 12 所述的装置, 其中, 所述掩码模块, 具体用于采用所述随机掩码对待部署模型的全部模型权重进行掩码处理, 得到第一噪声模型。

14. 根据权利要求 12 所述的装置, 其中, 所述目标参数还包括: 由第一字符和第二字符构建的筛选掩码, 其中, 所述筛选掩码的每一字符对应所述待部署模型的一个模型权重;

30 所述掩码模块, 具体用于采用所述随机掩码对所述待部署模型的目标模型权重进行掩码处理, 得到第一噪声模型;

其中, 所述目标模型权重为所述待部署模型中的对应所述第一字符的模型权重。

15. 根据权利要求 12 所述的装置, 其中, 所述目标参数还包括: 第三随机数种子, 所述掩码模块, 具体用于采用所述随机掩码对第二数量的目标模型层的模型权重进行掩码处理, 得到第一噪声模型;

35

其中，每一所述目标模型层在所述待部署模型中的层数与第二集合中的一个元素相等，所述第二集合包括基于所述第三随机数种子生成的第二数量的随机数。

16. 根据权利要求 12 所述的装置，其中，所述目标参数还包括：所述第二噪声模型的数字签名。

5 17. 根据权利要求 12 所述的装置，其中，所述发送模块，包括：

第一发送单元，用于将所述第二噪声模型发送至所述第二电子设备；

第二发送单元，用于在接收到所述第二电子设备发送的目标请求的情况下，将所述目标参数发送至所述第二电子设备；

10 其中，所述目标请求为所述第二电子设备在本地处于可信执行环境下发送的用于请求目标密钥的请求，所述目标密钥包括所述目标参数。

18. 一种模型部署的装置，应用于第二电子设备，所述装置包括：

接收模块，用于接收第一电子设备发送的目标参数和第二噪声模型，其中，所述目标参数包括第一随机数种子和第二随机数种子；所述第二噪声模型为第一噪声模型的多个模型层与第一数量的混淆层进行拼接得到的模型；所述第一噪声模型为采用随机掩码  
15 对待部署模型进行掩码处理得到的模型，所述随机掩码为基于所述第一随机数种子生成的掩码；所述混淆层为预设模型库中的模型层，每一所述混淆层在所述第二噪声模型中的层数与第一集合中的一个元素相等，所述第一集合包括基于所述第二随机数种子生成的第一数量的随机数；

20 第一去噪模块，用于基于所述目标参数去除所述第二噪声模型中的所述第一数量的混淆层，得到所述第一噪声模型；

第二去噪模块，用于基于所述目标参数去除所述第一噪声模型中的所述随机掩码，得到所述待部署模型。

25 19. 根据权利要求 18 所述的装置，其中，所述目标参数还包括：由第一字符和第二字符构建的筛选掩码，其中，所述筛选掩码的每一字符对应所述待部署模型的一个模型权重；

所述第二去噪模块，包括：

筛选单元，用于基于所述筛选掩码，确定所述第一噪声模型中的第一噪声模型权重，其中，所述第一噪声模型权重在所述第一噪声模型中的位置与目标模型权重在所述待部署模型中的位置相同，所述目标模型权重为所述待部署模型中的对应所述第一字符的模型  
30 权重；

第一随机数种子单元，用于基于所述第一随机数种子，生成所述随机掩码；

第一去噪单元，用于基于所述随机掩码对所述第一噪声模型中的所述第一噪声模型权重进行去掩码处理，得到所述待部署模型。

35 20. 根据权利要求 18 所述的装置，其中，所述目标参数还包括：第三随机数种子；  
所述第二去噪模块，包括：

第二随机数种子单元，用于基于所述第三随机数种子，生成第二数量的随机数；

确定单元，用于基于所述第二数量的随机数，确定所述第一噪声模型中第二数量的噪声模型层；其中，每一所述噪声模型层在所述第一噪声模型中的层数与所述第二数量的随机数中的一个随机数相等；

5 第三随机数种子单元，用于基于所述第一随机数种子，生成所述随机掩码；

第二去噪单元，用于基于所述随机掩码对第二噪声模型权重进行去掩码处理，得到所述待部署模型；其中，所述第二噪声模型权重为所述第二数量的噪声模型层的模型权重。

10 21. 根据权利要求 18 所述的装置，其中，所述目标参数还包括：所述第二噪声模型的数字签名；

所述第一去噪模块，具体用于在所述数字签名通过验签的情况下，基于所述目标参数去除所述第二噪声模型中的所述第一数量的混淆层，得到所述第一噪声模型。

22. 根据权利要求 18 所述的装置，其中，所述接收模块，具体用于：

接收第一电子设备发送的所述第二噪声模型；

15 在所述第二电子设备本地处于可信执行环境下，发送用于请求目标密钥的第一请求，其中，所述目标密钥包括所述目标参数；

接收所述第一电子设备基于所述第一请求返回的所述目标密钥。

20 23. 一种电子设备，包括处理器和存储器，所述存储器存储可在所述处理器上运行的程序或指令，所述程序或指令被所述处理器执行时实现如权利要求 1-6 任一项所述的模型部署的方法或如权利要求 7-11 任一项所述的模型部署的方法的步骤。

24. 一种可读存储介质，所述可读存储介质上存储程序或指令，所述程序或指令被处理器执行时实现如权利要求 1-6 任一项所述的模型部署的方法或如权利要求 7-11 任一项所述的模型部署的方法的步骤。

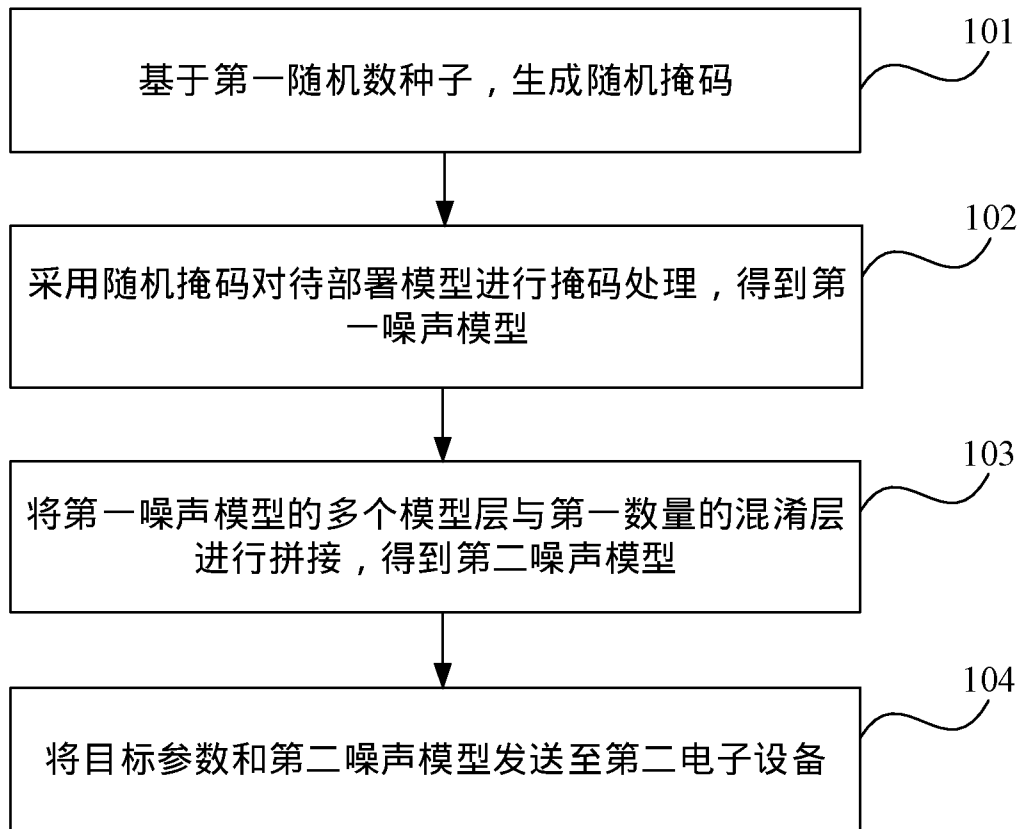


图 1

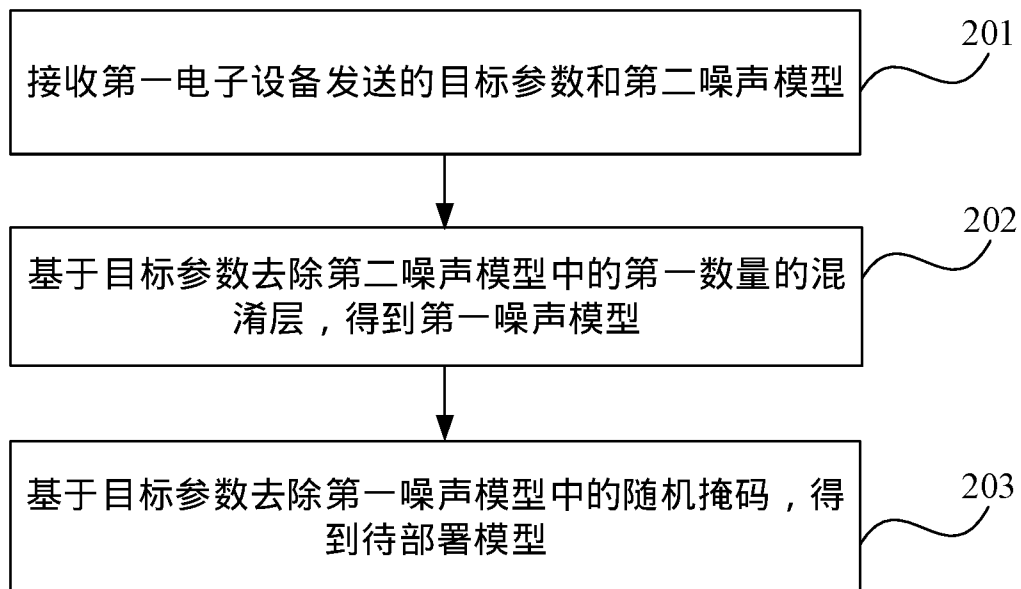


图 2

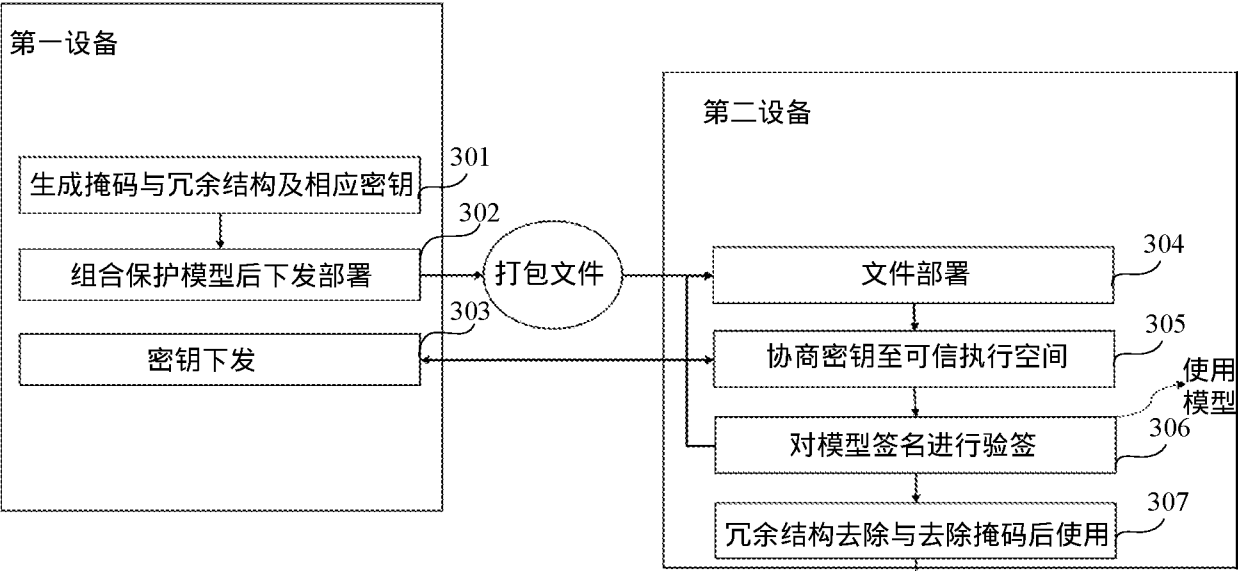


图 3

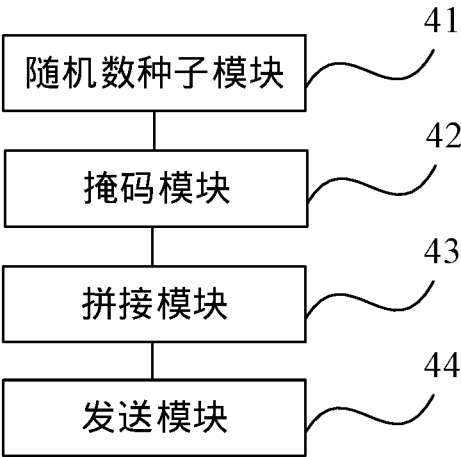


图 4

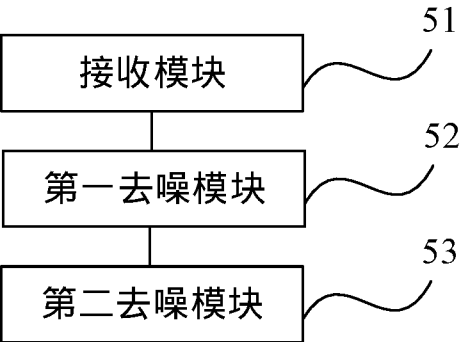


图 5

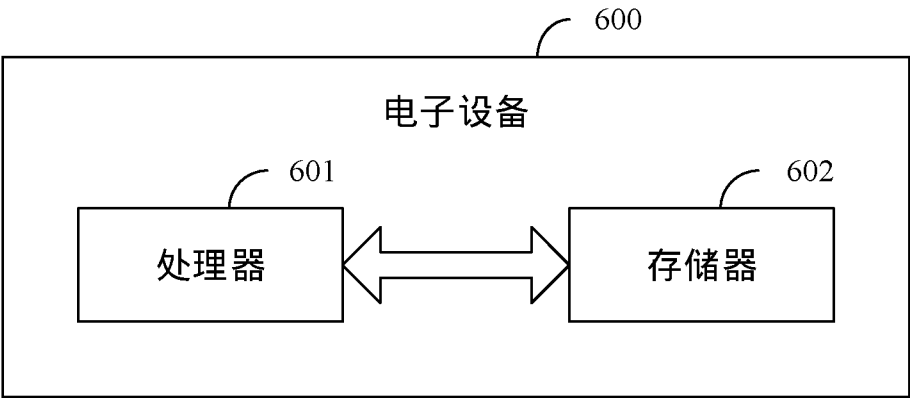


图 6

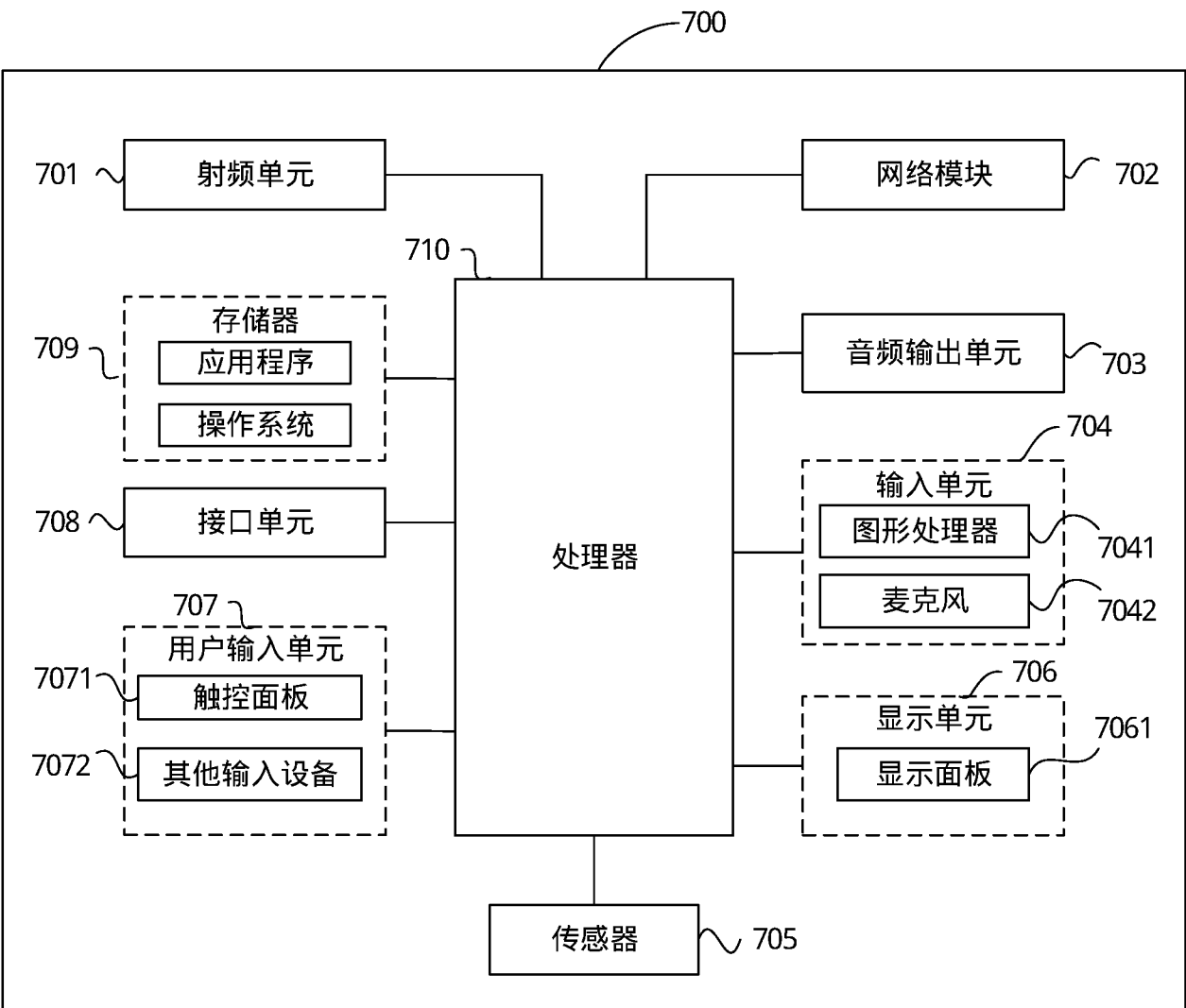


图 7

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2024/071178

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |                                                                                                                       |                                                                            |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------|
| <b>A. CLASSIFICATION OF SUBJECT MATTER</b><br>G06F21/14(2013.01)i<br><br>According to International Patent Classification (IPC) or to both national classification and IPC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                                                                                                                       |                                                                            |
| <b>B. FIELDS SEARCHED</b><br><br>Minimum documentation searched (classification system followed by classification symbols)<br>IPC:G06F<br><br>Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched<br><br>Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)<br>CNABS; CNTXT; DWPI; SIPOABS; USTXT; CNKI; IEEE: 模型, 随机数, 种子, 掩码, 噪声, 混淆, 添加, 增加, 混合, 拼接, 层, model, random number, seed, mask, noise, confusion, add, mix, splice, layer                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |                                                                                                                       |                                                                            |
| <b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                                                                                                                       |                                                                            |
| Category*                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | Citation of document, with indication, where appropriate, of the relevant passages                                    | Relevant to claim No.                                                      |
| PX                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | CN 116010911 A (VIVO COMMUNICATION TECHNOLOGY CO., LTD.) 25 April 2023 (2023-04-25)<br>description, paragraphs 41-120 | 1-24                                                                       |
| A                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | CN 112749392 A (XIDIAN UNIVERSITY) 04 May 2021 (2021-05-04)<br>description, paragraphs 17-23, 54, and 72              | 1-24                                                                       |
| A                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | CN 115238827 A (ALIPAY (HANGZHOU) INFORMATION TECHNOLOGY CO., LTD.) 25 October 2022 (2022-10-25)<br>entire document   | 1-24                                                                       |
| A                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | CN 115168841 A (HUAWEI TECHNOLOGIES CO., LTD.) 11 October 2022 (2022-10-11)<br>entire document                        | 1-24                                                                       |
| A                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | US 2021406589 A1 (AMAZON TECHNOLOGIES, INC.) 30 December 2021 (2021-12-30)<br>entire document                         | 1-24                                                                       |
| A                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | US 2021256309 A1 (XAIN AG) 19 August 2021 (2021-08-19)<br>entire document                                             | 1-24                                                                       |
| <input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                                                                       |                                                                            |
| <p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“D” document cited by the applicant in the international application</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p> <p>“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>“&amp;” document member of the same patent family</p> |                                                                                                                       |                                                                            |
| Date of the actual completion of the international search<br><b>25 March 2024</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                                                                                                                       | Date of mailing of the international search report<br><b>10 April 2024</b> |
| Name and mailing address of the ISA/CN<br><b>China National Intellectual Property Administration (ISA/CN)<br/>China No. 6, Xitucheng Road, Jimenqiao, Haidian District,<br/>Beijing 100088</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                                                                                                       | Authorized officer<br><br><br><br>Telephone No.                            |

INTERNATIONAL SEARCH REPORT  
Information on patent family members

International application No.  
**PCT/CN2024/071178**

| Patent document cited in search report |            |    | Publication date (day/month/year) | Patent family member(s) |            |    | Publication date (day/month/year) |
|----------------------------------------|------------|----|-----------------------------------|-------------------------|------------|----|-----------------------------------|
| CN                                     | 116010911  | A  | 25 April 2023                     | None                    |            |    |                                   |
| CN                                     | 112749392  | A  | 04 May 2021                       | CN                      | 112749392  | B  | 04 October 2022                   |
| CN                                     | 115238827  | A  | 25 October 2022                   | CN                      | 115238827  | B  | 25 November 2022                  |
| CN                                     | 115168841  | A  | 11 October 2022                   | None                    |            |    |                                   |
| US                                     | 2021406589 | A1 | 30 December 2021                  | WO                      | 2021262399 | A1 | 30 December 2021                  |
|                                        |            |    |                                   | US                      | 11334773   | B2 | 17 May 2022                       |
|                                        |            |    |                                   | US                      | 2022405528 | A1 | 22 December 2022                  |
|                                        |            |    |                                   | US                      | 11854116   | B2 | 26 December 2023                  |
| US                                     | 2021256309 | A1 | 19 August 2021                    | US                      | 11487969   | B2 | 01 November 2022                  |

A. 主题的分类

G06F21/14(2013.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

IPC:G06F

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNABS;CNTXT;DWPI;SIPOABS;USTXT;CNKI;IEEE;模型, 随机数, 种子, 掩码, 噪声, 混淆, 添加, 增加, 混合, 拼接, 层, model, random number, seed, mask, noise, confusion, add, mix, splice, layer

C. 相关文件

| 类型* | 引用文件, 必要时, 指明相关段落                                                        | 相关的权利要求 |
|-----|--------------------------------------------------------------------------|---------|
| PX  | CN 116010911 A (维沃移动通信有限公司) 2023年4月25日 (2023 - 04 - 25)<br>说明书第41-120段   | 1-24    |
| A   | CN 112749392 A (西安电子科技大学) 2021年5月4日 (2021 - 05 - 04)<br>说明书第17-23、54、72段 | 1-24    |
| A   | CN 115238827 A (支付宝(杭州)信息技术有限公司) 2022年10月25日 (2022 - 10 - 25)<br>全文      | 1-24    |
| A   | CN 115168841 A (华为技术有限公司) 2022年10月11日 (2022 - 10 - 11)<br>全文             | 1-24    |
| A   | US 2021406589 A1 (AMAZON TECH INC) 2021年12月30日 (2021 - 12 - 30)<br>全文    | 1-24    |
| A   | US 2021256309 A1 (XAIN AG) 2021年8月19日 (2021 - 08 - 19)<br>全文             | 1-24    |

☐ 其余文件在C栏的续页中列出。

☒ 见同族专利附件。

\* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“D” 申请人在国际申请中引证的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“p” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2024年3月25日

国际检索报告邮寄日期

2024年4月10日

ISA/CN的名称和邮寄地址

中国国家知识产权局  
中国北京市海淀区蓟门桥西土城路6号 100088

授权官员

刘念

电话号码 (+86) 027-59371893

PCT/ISA/210 表(第2页) (2022年7月)

国际检索报告  
关于同族专利的信息

国际申请号  
PCT/CN2024/071178

| 检索报告引用的专利文件 |            |    | 公布日<br>(年/月/日) | 同族专利 | 公布日<br>(年/月/日)            |
|-------------|------------|----|----------------|------|---------------------------|
| CN          | 116010911  | A  | 2023年4月25日     | 无    |                           |
| CN          | 112749392  | A  | 2021年5月4日      | CN   | 112749392 B 2022年10月4日    |
| CN          | 115238827  | A  | 2022年10月25日    | CN   | 115238827 B 2022年11月25日   |
| CN          | 115168841  | A  | 2022年10月11日    | 无    |                           |
| US          | 2021406589 | A1 | 2021年12月30日    | WO   | 2021262399 A1 2021年12月30日 |
|             |            |    |                | US   | 11334773 B2 2022年5月17日    |
|             |            |    |                | US   | 2022405528 A1 2022年12月22日 |
|             |            |    |                | US   | 11854116 B2 2023年12月26日   |
| US          | 2021256309 | A1 | 2021年8月19日     | US   | 11487969 B2 2022年11月1日    |