(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 2 月 15 日 (15.02.2024)

WIPO | PCT

(10) 国际公布号
WO 2024/031886 A1

(51) 国际专利分类号:
H04L 9/00 (2022.01) *H04L 9/08* (2006.01)
G06F 21/60 (2013.01)

(21) 国际申请号: PCT/CN2022/137361

(22) 国际申请日: 2022 年 12 月 7 日 (07.12.2022)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202210952494.8 2022年8月9日 (09.08.2022) CN

(71) 申请人: 中国银联股份有限公司 (**CHINA UNIONPAY CO., LTD.**) [CN/CN]; 中国上海市浦东新区含笑路36号银联大厦, Shanghai 200135 (CN)。

(72) 发明人: 何朔 (**HE, Shuo**); 中国上海市浦东新区含笑路36号银联大厦, Shanghai 200135 (CN)。 刘红宝 (**LIU, Hongbao**); 中国上海市浦东新区含笑路36号银联大厦, Shanghai 200135 (CN)。 郑建宾 (**ZHENG, Jianbin**); 中国上海市浦东新区含笑路36号银联大厦, Shanghai 200135 (CN)。 高鹏飞 (**GAO, Pengfei**); 中国上海市浦东新区含笑路36号银联大厦, Shanghai 200135 (CN)。 李亮 (**LI, Liang**); 中国上海市浦东新区含笑路36号银联大厦, Shanghai 200135 (CN)。

(74) 代理人: 北京同达信恒知识产权代理有限公司 (**TDIP & PARTNERS**); 中国北京市西城区裕民路18号北环中心A座2002, Beijing 100029 (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, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE,

(54) **Title:** DATA MATCHING METHOD, APPARATUS AND SYSTEM, AND DEVICE AND MEDIUM

(54) 发明名称: 一种数据匹配方法、装置、系统、设备及介质

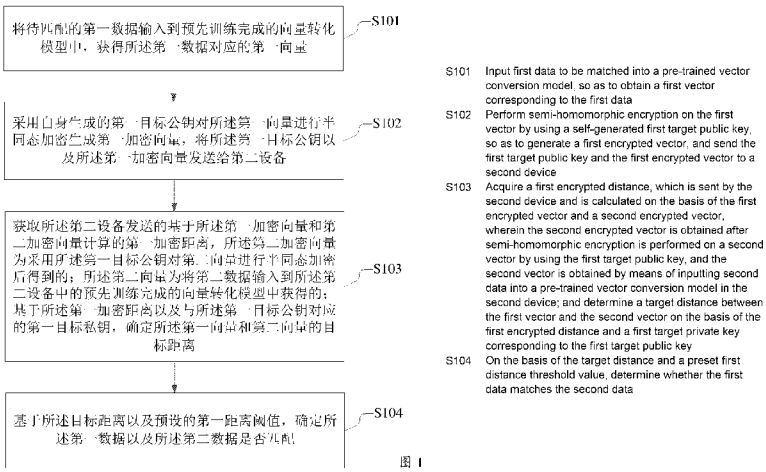


图 1

(57) **Abstract:** Disclosed in the present application are a data matching method, apparatus and system, and a device and a medium, which are used for solving the problem of a secure intersection algorithm in the prior art being able to perform secure intersection only when data of both parties is completely the same, resulting in a limited number of usage scenarios and affecting the service range of data matching. In the present application, a first vector corresponding to first data and a second vector corresponding to second data are respectively obtained; a first encrypted vector obtained by means of encrypting the first vector and a second encrypted vector obtained by means of encrypting the second vector are acquired; a first encrypted distance is calculated on the basis of the first encrypted vector and the second encrypted vector; a target distance between the first vector and the second vector is determined on the basis of the first encrypted distance and a first target private key; and on the basis of the target distance and a preset first distance threshold value, it is determined whether the first data matches the second data. In this way, fuzzy matching between the first data and the second data can be realized, such that usage scenarios are widened, and the security of a matching process is ensured.

WO 2024/031886 A1

[见续页]

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: 本申请公开了一种数据匹配方法、装置、系统、设备及介质，用以解决现有技术中的安全求交算法只能在双方数据完全相同时进行安全求交，使用场景有限，影响数据匹配的业务范围的问题。由于本申请可以分别获得该第一数据对应的第一向量以及第二数据对应的第二向量，并获取基于该第一向量加密后的第一加密向量，以及该第二向量加密后的第二加密向量，基于第一加密向量和第二加密向量计算第一加密距离，并基于第一加密距离以及第一目标私钥，确定第一向量和第二向量的目标距离，基于该目标距离以及预设的第一距离阈值确定第一数据和第二数据是否匹配，可以实现第一数据和第二数据的模糊匹配，拓宽了使用场景，保证了匹配过程的安全性。

一种数据匹配方法、装置、系统、设备及介质

相关申请的交叉引用

本申请要求在2022年08月09日提交中国专利局、申请号为202210952494.8、申请名称为“一种数据匹配方法、装置、系统、设备及介质”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

本申请涉及数据处理技术领域，尤其涉及一种数据匹配方法、装置、系统、设备及介质。

背景技术

当前隐私计算技术主要应用于安全求交和联邦学习。所谓安全求交是指识别双方数据的交集，比如，即识别机构 A 和机构 B 的共有用户，其中安全求交也为纵向联邦学习的第一个步骤，也就是说，先对手机号、身份证号、营业执照号等关键信息进行安全求交，然后再进行下一步联合建模等步骤。

相关技术中，为了识别双方数据的交集或者实现双方数据的匹配，常见的安全求交算法包括基于 RSA 加密算法的安全求交算法等。但是目前的安全求交算法只有当双方数据完全相同时，也就是说，当双方数据的数据类型以及数据包含的字符的数量完全相同时，才能匹配成功。但是实际业务中，往往存在很多数据不完全相同时待匹配的使用场景，因此现有技术中的安全求交算法极大限制了其使用场景，影响匹配的业务范围。

发明内容

本申请提供了一种数据匹配方法、装置、系统、设备及介质，用以解决现有技术中的安全求交算法只能在双方数据完全相同时进行安全求交，使用场景有限，影响数据匹配的业务范围的问题。

第一方面，本申请提供了一种数据匹配方法，应用于第一设备，所述方法包括：

将待匹配的第一数据输入到预先训练完成的向量转化模型中，获得所述第一数据对应的第一向量；

采用自身生成的第一目标公钥对所述第一向量进行半同态加密生成第一加密向量，将所述第一目标公钥以及所述第一加密向量发送给第二设备；

获取所述第二设备发送的基于所述第一加密向量和第二加密向量计算的第一加密距离，所述第二加密向量为采用所述第一目标公钥对第二向量进行半同态加密后得到的；所述第二向量为将第二数据输入到所述第二设备中的预先训练完成的向量转化模型中获得的；

基于所述第一加密距离以及与所述第一目标公钥对应的第一目标私钥，确定所述第一向量和第二向量的目标距离；

基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配。

第二方面，本申请提供了一种数据匹配方法，应用于第二设备，所述方法包括：

将待匹配的第二数据输入到预先训练完成的向量转化模型中，获得所述第二数据对应的第二向量；

接收第一设备发送的第一目标公钥以及第一加密向量，采用所述第一目标公钥对所述第二向量进行半同态加密生成第二加密向量；其中，所述第一加密向量为采用所述第一目标公钥对第一向量进行半同态加密后得到的，所述第一向量为将第一数据输入到所述第一设备中的预先训练完成的向量转化模型中获得的；

基于所述第一加密向量和第二加密向量计算第一加密距离，将所述第一加密距离发送给所述第一设备，以使所述第一设备基于所述第一加密距离以及与所述第一目标公钥对应的第一目标私钥，确定所述第一向量和第二向量的目标距离；并使所述第一设备基于所述目标距离以及预设的第一距离阈值，

确定所述第一数据以及所述第二数据是否匹配。

第三方面，本申请提供了一种数据匹配装置，应用于第一设备，所述装置包括:

第一获取模块，用于将待匹配的第一数据输入到预先训练完成的向量转化模型中，获得所述第一数据对应的第一向量;

第一处理模块，用于采用自身生成的第一目标公钥对所述第一向量进行半同态加密生成第一加密向量，将所述第一目标公钥以及所述第一加密向量发送给第二设备;

所述第一获取模块，还用于获取所述第二设备发送的基于所述第一加密向量和第二加密向量计算的第一加密距离，所述第二加密向量为采用所述第一目标公钥对第二向量进行半同态加密后得到的；所述第二向量为将第二数据输入到所述第二设备中的预先训练完成的向量转化模型中获得的;

第一确定模块，用于基于所述第一加密距离以及与所述第一目标公钥对应的第一目标私钥，确定所述第一向量和第二向量的目标距离；基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配。

第四方面，本申请提供了一种数据匹配装置，应用于第二设备，所述装置包括:

第二获取模块，用于将待匹配的第二数据输入到预先训练完成的向量转化模型中，获得所述第二数据对应的第二向量;

第二处理模块，用于接收第一设备发送的第一目标公钥以及第一加密向量，采用所述第一目标公钥对所述第二向量进行半同态加密生成第二加密向量；其中，所述第一加密向量为采用所述第一目标公钥对第一向量进行半同态加密后得到的，所述第一向量为将第一数据输入到所述第一设备中的预先训练完成的向量转化模型中获得的;

第二确定模块，用于基于所述第一加密向量和第二加密向量计算第一加密距离，将所述第一加密距离发送给所述第一设备，以使所述第一设备基于

所述第一加密距离以及与所述第一目标公钥对应的第一目标私钥，确定所述第一向量和第二向量的目标距离；并使所述第一设备基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配。

第五方面，本申请提供了一种数据匹配系统，所述系统包括：

第一设备，用于将待匹配的第一数据输入到预先训练完成的向量转化模型中，获得所述第一数据对应的第一向量；采用自身生成的第一目标公钥对所述第一向量进行半同态加密生成第一加密向量，将所述第一目标公钥以及所述第一加密向量发送给第二设备；

所述第二设备，用于将待匹配的第二数据输入到预先训练完成的向量转化模型中，获得所述第二数据对应的第二向量；并接收所述第一设备发送的第一目标公钥以及第一加密向量，采用所述第一目标公钥对所述第二向量进行半同态加密生成第二加密向量；基于所述第一加密向量和第二加密向量计算第一加密距离，将所述第一加密距离发送给所述第一设备；

所述第一设备，还用于基于所述第一加密距离以及与所述第一目标公钥对应的第一目标私钥，确定所述第一向量和第二向量的目标距离；基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配。

第六方面，本申请提供了一种电子设备，所述电子设备包括处理器和存储器，所述存储器用于存储程序指令，所述处理器用于执行存储器中存储的计算机程序时实现上述任一所述数据匹配方法的步骤。

第七方面，本申请提供了一种计算机可读存储介质，其存储有计算机程序，所述计算机程序被处理器执行时实现上述任一所述数据匹配方法的步骤。

由于在本申请实施例中，可以分别将待匹配的第一数据和第二数据输入到预先训练完成的向量转化模型中，获得该第一数据对应的第一向量以及第二数据对应的第二向量，并获取基于该第一向量加密后的第一加密向量，以及该第二向量加密后的第二加密向量，基于第一加密向量和第二加密向量计算第一加密距离，并基于第一加密距离以及第一目标私钥，确定第一向量和

第二向量的目标距离，基于该目标距离以及预设的第一距离阈值确定第一数据和第二数据是否匹配，即可以在第一数据和第二数据不完全相同时，也能实现第一数据和第二数据的模糊匹配，拓宽了使用场景，且在进行模糊匹配过程中引入了第一目标公钥和第一目标私钥分别进行半同态加密和解密，实现了安全求交，保证了匹配过程的安全性，且整个匹配的过程中，第一数据以及第二数据均未以原始数据的形式离开过对应的第一设备以及第二设备，保证了第一数据和第二数据的安全性，实现了原始数据不出库也能实现模糊匹配，进一步保证了匹配过程的安全性。

附图说明

为了更清楚地说明本发明实施例中的技术方案，下面将对实施例描述中所需要使用的附图作简要介绍，显而易见地，下面描述中的附图仅仅是本发明的一些实施例，对于本领域的普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

图 1 为本申请实施例提供的一种数据匹配过程示意图；

图 2 为本申请实施例提供的第二种数据匹配过程示意图；

图 3 为本申请实施例提供的第三种数据匹配过程示意图；

图 4 为本申请一些实施例提供的一种数据匹配装置结构示意图；

图 5 为本申请一些实施例提供的另一种数据匹配装置结构示意图；

图 6 为本申请一些实施例提供的一种数据匹配系统结构示意图；

图 7 为本申请一些实施例提供的一种电子设备结构示意图。

具体实施方式

为了保证双方数据不完全相同时也可以进行匹配，拓宽数据匹配的业务范围，本申请实施例提供了一种数据匹配方法、装置、系统、设备及介质。

为使本申请的目的和实施方式更加清楚，下面将结合本申请示例性实施例中的附图，对本申请示例性实施方式进行清楚、完整地描述，显然，描述

的示例性实施例仅是本申请一部分实施例，而不是全部的实施例。

需要说明的是，本申请中对于术语的简要说明，仅是为了方便理解接下来描述的实施方式，而不是意图限定本申请的实施方式。除非另有说明，这些术语应当按照其普通和通常的含义理解。

本申请中说明书和权利要求书及上述附图中的术语“第一”、“第二”、“第三”等是用于区别类似或同类的对象或实体，而不必然意味着限定特定的顺序或先后次序，除非另外注明。应该理解这样使用的用语在适当情况下可以互换。

术语“包括”和“具有”以及他们的任何变形，意图在于覆盖但不排他的包含，例如，包含了一系列组件的产品或设备不必限于清楚地列出的所有组件，而是可包括没有清楚地列出的或对于这些产品或设备固有的其它组件。

术语“模块”是指任何已知或后来开发的硬件、软件、固件、人工智能、模糊逻辑或硬件或/和软件代码的组合，能够执行与该元件相关的功能。

最后应说明的是：以上各实施例仅用以说明本申请的技术方案，而非对其限制；尽管参照前述各实施例对本申请进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分或者全部技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本申请各实施例技术方案的范围。

实施例 1:

图 1 为本申请实施例提供的第一种数据匹配过程示意图，该过程包括以下步骤:

S101: 将待匹配的第一数据输入到预先训练完成的向量转化模型中，获得所述第一数据对应的第一向量。

本申请实施例提供的数据匹配方法应用于电子设备（为方便描述，称为第一设备），第一设备可以是智能终端、PC 或者服务器等设备。

为了保证双方数据不完全相同时也能实现模糊匹配，在本申请实施例中，第一设备中部署了预先训练完成的向量转化模型，该预先训练完成的向量转

化模型用于获得待匹配的数据对应的向量，且针对不同的数据，该预先训练完成的向量转化模型输出的向量的维度可以相同。

在一种可能的实施方式中，为了获得待匹配的第一数据（为方便描述，将第一设备中保存的待匹配的数据称为第一数据）对应的第一向量（为方便描述，将第一数据对应的向量称为第一向量），可以将该第一数据输入到预先训练完成的向量转化模型中，该预先训练完成的向量转化模型可以输出该第一数据对应的第一向量，可选的，该第一向量中的每个分量可以为数字，即可以通过该预先训练完成的向量转化模型将第一数据进行了量化。示例性的，以第一数据为“某海市某东新区晴某小卖部”为例，可以将该第一数据“某海市某东新区晴某小卖部”输入到预先训练完成的词向量模型中，输出的该第一数据“某海市某东新区晴某小卖部”对应的第一向量可以为（1.0，2.0，1.5，2.0，3.5）。

S102：采用自身生成的第一目标公钥对所述第一向量进行半同态加密生成第一加密向量，将所述第一目标公钥以及所述第一加密向量发送给第二设备。

在一种可能的实施方式中，为了提高安全性，第一设备可以生成第一目标公私钥对，其中，该第一目标公私钥对中包含第一目标公钥和第一目标私钥（为方便描述，将第一设备生成的公钥称为第一目标公钥，将第一设备生成的私钥称为第一目标私钥）。可选的，第一目标公私钥对可以为半同态加密公私钥对，即第一目标公钥可以为半同态加密公钥，第一目标私钥可以为半同态加密私钥。第一设备可以根据自身生成的第一目标公钥对第一向量进行半同态加密生成第一加密向量。其中，该第一目标公私钥对可以为对称公私钥对，也可以为非对称公私钥对，可以根据需求设置目标公私钥对。其中，生成第一目标公私钥对的过程为现有技术，在此不做赘述。

在一种可能的实施方式中，由于待与第一数据进行匹配的第二数据为第二设备获取的，为了便于后续能够确定第一数据对应的第一向量和第二数据对应的第二向量之间的目标距离，第一设备可以将第一目标公钥以及第一加

密向量发送给第二设备，第二设备可以根据该第一目标公钥，对第二数据对应的第二向量进行半同态加密生成第二加密向量。

S103：获取所述第二设备发送的基于所述第一加密向量和第二加密向量计算的第一加密距离，所述第二加密向量为采用所述第一目标公钥对第二向量进行半同态加密后得到的；所述第二向量为将第二数据输入到所述第二设备中的预先训练完成的向量转化模型中获得的；基于所述第一加密距离以及与所述第一目标公钥对应的第一目标私钥，确定所述第一向量和第二向量的目标距离。

在一种可能的实施方式中，为了实现第一数据和第二数据的模糊匹配，第二设备中也部署了预先训练完成的向量转化模型，用于获得待匹配的第二数据（为方便描述，将第二设备中待匹配的数据称为第二数据）对应的第二向量（为方便描述，将第二数据对应的向量称为第二向量）。可选的，可以将第二数据输入到预先训练完成的向量转化模型中，该预先训练完成的向量转化模型可以输出该第二数据对应的第二向量，第二设备可以基于接收到的第一设备发送的第一目标公钥对该第二向量进行半同态加密，获得第二加密向量。

为了确定第一向量和第二向量之间的距离，在一种可能的实施方式中，第二设备可以针对第一加密向量中的每个第一加密分量、第二向量中的每个第二分量以及第二加密向量中的每个第二加密分量，通过半同态加密算法，确定该第一加密分量的第一加密平方分量、加密后的该第一加密分量与相应的第二分量的乘积，以及相应的第二加密分量的第二加密平方分量；基于该第一加密平方分量、该乘积以及该第二加密平方分量，确定该第一加密分量对应的加密后的子距离；并可以基于每个第一加密分量对应的加密后的子距离，确定第一加密向量和第二加密向量的第一加密距离。为方便理解，下面以公式形式对本申请实施例提供的确定第一加密向量和第二加密向量的第一加密距离的过程进行解释说明。

示例性的，假设第一数据 U1 对应的第一向量为（x_1，x_2，x_3……，x_m），

第二数据U5对应的第二向量为（y_1，y，y_3……，y_m）。

第一设备生成第一目标公私钥对A (pka1，ska1)，其中，该pka1为第一目标公钥，ska1为第一目标私钥，基于该第一目标公钥对第一向量进行半同态加密生成第一加密向量，示例性的，针对第一向量中的每个第一分量，可以基于第一目标公钥对第一向量中的每个第一分量分别进行半同态加密，生成第一加密向量。例如第一向量（x_1，x_2，x_3……，x_m）对应的第一加密向量为（$E_{pka1}(x_1)$，$E_{pka1}(x_2)$，$E_{pka1}(x_3)$……，$E_{pka1}(x_m)$）。第一加密向量中包含m个第一加密分量$E_{pka1}(x_i)$，其中，i为不大于m的任一正整数。第一设备可以将该第一加密向量以及该第一目标公钥发送给第二设备。

第二设备接收到第一设备发送的第一目标公钥以及第一加密向量后，可以基于该第一目标公钥对第二向量进行半同态加密，获得第二加密向量。示例性的，针对第二向量中的每个第二分量，可以基于第一目标公钥对第二向量中的每个第二分量分别进行半同态加密，生成第二加密向量。例如第二向量（y_1，y_2，y_3……，y_m）对应的第二加密向量为（$E_{pka1}(y_1)$，$E_{pka1}(y_2)$，$E_{pka1}(y_3)$……，$E_{pka1}(y_m)$），第二向量中包含m个第二分量y_i，第二加密向量中包含m个第二加密分量$E_{pka1}(y_i)$，其中，i为不大于m的任一正整数。

在一种可能的实施方式中，针对第一加密向量中的每个第一加密分量$E_{pka1}(x_i)$，第二设备可以确定该第一加密分量的第一加密平方分量：$E_{pka1}(x_i^2)$。另外，针对第一加密向量中的每个第一加密分量$E_{pka1}(x_i)$，第二设备还可以通过半同态加密算法，确定加密后的该第一加密分量与相应的第二加密分量的乘积，另外，第二设备还可以确定相应的第二加密分量的第二加密平方分量：$E_{pka1}(y_i^2)$。

在一种可能的实施方式中，通过半同态加密算法，确定加密后的第一加密分量与相应的第二加密分量的乘积的过程可以如下：

通过半同态加密算法，确定加密后的指数幂，其中，该指数幂的底数可以为该第一加密分量，该指数幂的指数可以为相应的第二分量。例如，指数

幂可以为$E_{pka1}(x_i)^{y_i}$，该指数幂可以用于确定基于半同态加密算法加密后的该第一加密分量与相应的第二分量的乘积。其中，针对如何基于该指数幂确定第一加密分量与相应的第二分量的乘积（$E_{pka1}(x_iy_i)$）在后续进行详细介绍，在此先不赘述。

在一种可能的实施方式中，针对每个第一加密分量，可以基于该第一加密分量对应的第一加密平方分量$E_{pka1}(x_i^2)$、加密后的该第一加密分量与相应的第二加密分量的乘积$E_{pka1}(x_i)^{y_i}$以及第二加密平方分量$E_{pka1}(y_i^2)$，确定该第一加密分量对应的加密后的子距离：$E_{pka1}(x_i^2)$-$2E_{pka1}(x_i)^{y_i}$+$E_{pka1}(y_i^2)$。可选的，可以将每个第一加密分量对应的加密后的子距离的和值，确定为第一加密向量和第二加密向量的第一加密距离：$\sum_{i=1}^{m}(E_{pka1}(x_i^2)-2E_{pka1}(x_i)^{y_i}+E_{pka1}(y_i^2))$。

在一种可能的实施方式中，第二设备获得第一加密距离之后，可以将该第一加密距离发送给第一设备。第一设备接收到该第一加密距离之后，可以根据第一设备生成的第一目标私钥对中的第一目标私钥，对该第一加密距离进行解密，确定第一向量和第二向量的目标距离。

示例性的，第一设备可以根据第一目标私钥，对第一加密距离进行解密后，得到$\sum_{i=1}^{m}(x_i^2-2x_iy_i+y_i^2)$。

为方便理解，下面针对第一设备如何基于指数幂$E_{pka1}(x_i)^{y_i}$，确定加密后的第一分量与相应的第二分量的乘积（$E_{pka1}(x_iy_i)$）的过程进行解释说明。

在半同态加密算法中，可以将对$E_{pka1}(x_i)$的运算看做为对第一指数幂$g^{E_{pka1}(x_i)}$的运算，即$E_{pka1}(x_i)=g^{E_{pka1}(x_i)}$，其中 g 可以为任意底数。

由于$g^{x+y}=g^x*g^y$，则$E_{pka1}(x_i+y_i)=E_{pka1}(x_i)*E_{pka1}(y_i)$。

又由于x_iy_i等于y_i个x_i相加，因此，

$E_{pka1}(x_iy_i)=E_{pka1}(x_i+x_i+\cdots+x_i$，共$y_i$个$x_i$相加）

$=g^{E_{pka1}(x_i)}*g^{E_{pka1}(x_i)}*\ldots\ldots*g^{E_{pka1}(x_i)}$（共$y_i$个$g^{E_{pka1}(x_i)}$相乘）

$=\left(g^{E_{pka1}(x_i)}\right)^{y_i}$

$=E_{pka1}(x_i)^{y_i}$。

由上述公式可以看出，通过半同态加密算法，可以基于指数幂$E_{pka1}(x_i)^{y_i}$，确定加密后的第一分量与相应的第二分量的乘积（$E_{pka1}(x_i y_i)$），第一设备进而可以根据第一目标私钥，对第一加密距离进行解密后，得到$\sum_{i=1}^{m}(x_i^2 - 2x_i y_i + y_i^2)$。第一设备可以基于解密后的第一加密距离以及欧氏距离公式，计算第一向量和第二向量之间的目标距离，示例性的，第一向量和第二向量之间的目标距离$d = \sqrt{\sum_{i=1}^{m}(x_i^2 - 2x_i y_i + y_i^2)}$。其中，该$x_i$为第一向量中的第 i 个分量，$y_i$为第二向量中的第 i 个分量，m 为第一向量或第二向量中包含的分量的数量，且该第一向量中包含的分量的数量与该第二分量中包含的分量的数量相同，即第一向量的长度与第二向量的长度相等，第一向量和第二向量的长度可以均为预设长度 m，m 可以为任意正整数。

在另外一种可能的实施方式中，还可以基于余弦距离公式或者汉明距离公式等确定第一向量与第二向量之间的距离，在此不再赘述。

另外，本申请实施例还提供了另一种确定第一加密距离的方式，在一种可能的实施方式中，第一加密距离还可以采用以下过程获得:

针对第一加密向量中的每个第一加密分量$E_{pka1}(x_i)$、第二向量中的每个第二分量y_i以及第二加密向量中的每个第二加密分量$E_{pka1}(y_i)$，第二设备可以通过半同态加密算法，确定加密后的该第一加密分量与相应的第二分量的乘积$E_{pka1}(x_i)^{y_i}$，以及相应的第二加密分量的第二加密平方分量$E_{pka1}(y_i^2)$；第二设备可以基于该乘积$E_{pka1}(x_i)^{y_i}$以及该第二加密平方分量$E_{pka1}(y_i^2)$，确定该第一加密分量对应的加密后的子距离：$b = E_{pka1}(-2E_{pka1}(x_i)^{y_i} + y_i^2)$。

其中，通过半同态加密算法，确定第一加密分量与相应的第二分量的乘积的过程与上述实施例相同，例如可以通过半同态加密算法，确定加密后的指数幂，其中，该指数幂的底数为该第一加密分量，该指数幂的指数为相应的第二分量；该指数幂用于确定基于半同态加密算法加密后的该第一加密分量与相应的第二分量的乘积，在此不再赘述。

第二设备可以将该加密后的子距离 b 发送给第一设备，第一设备针对第一加密向量中的每个第一加密分量$E_{pka1}(x_i)$，可以确定该第一加密分量的第一加密平方分量$E_{pka1}(x_i^2)$，并可以基于该第一加密平方分量$E_{pka1}(x_i^2)$以及加密后的子距离$E_{pka1}(-2E_{pka1}(x_i)^{y_i}+y_i^2)$，对该第一加密分量对应的加密后的子距离进行更新，可选的，可以将加密后的子距离更新为：

$E_{pka1}(x_i^2)$-$2E_{pka1}(x_i)^{y_i}$+$E_{pka1}(y_i^2)$。

第一设备可以基于每个第一加密分量对应的更新后的子距离，确定第一加密向量和第二加密向量的更新后的第一加密距离。可选的，可以将每个第一加密分量对应的更新后的子距离的和值，确定为第一加密向量和第二加密向量的更新后的第一加密距离：$\sum_{i=1}^{m}(E_{pka1}(x_i^2)-2E_{pka1}(x_i)^{y_i}+E_{pka1}(y_i^2))$。与上述实施例类似，确定了更新后的第一加密距离之后，第一设备可以根据第一设备生成的第一目标私钥对中的第一目标私钥，对该更新后的第一加密距离进行解密，以确定第一向量和第二向量的目标距离。

示例性的，第一设备可以根据第一目标私钥，对更新后的第一加密距离进行解密后，得到$\sum_{i=1}^{m}(x_i^2-2x_iy_i+y_i^2)$，可以基于解密后的第一加密距离以及欧氏距离公式，计算第一向量和第二向量之间的目标距离，示例性的，第一向量和第二向量之间的目标距离$d=\sqrt{\sum_{i=1}^{m}(x_i^2-2x_iy_i+y_i^2)}$，在此不再赘述。

S104：基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配。

在一种可能的实施方式中，为了确定第一数据与第二数据是否匹配，确定了第一向量与第二向量之间的目标距离后，可以将该目标距离与预设的第一距离阈值进行比较，根据比较结果，确定该第一数据与该第二数据是否匹配。示例性的，可以判断第一向量与第二向量之间的目标距离是否小于预设的第一距离阈值，如果第一向量与第二向量之间的目标距离小于预设的第一距离阈值，则可以认为第一数据与第二数据匹配。而如果第一向量与第二向

量之间的目标距离不小于预设的第一距离阈值，则可以确定第一数据与第二数据不匹配。可选的，预设的第一距离阈值可以为 1，还可以为 1.5 等，本申请对第一距离阈值不作具体限定，可以根据需求灵活设置。其中，第一向量与第二向量之间的目标距离越小，第一向量与第二向量越匹配。

为了确定第二数据与第一数据是否完全匹配，在一种可能的实施方式中，还可以判断第一向量与第二向量之间的目标距离是否等于预设的第二距离阈值，若该目标距离等于预设的第二距离阈值，则可以认为第一数据与第二数据相同，也就是说，该第一数据与该第二数据完全匹配。可选的，该第二距离阈值可以小于第一距离阈值，且第二距离阈值可以为 0。

由于在本申请实施例中，可以分别将待匹配的第一数据和第二数据输入到预先训练完成的向量转化模型中，获得该第一数据对应的第一向量以及第二数据对应的第二向量，并获取基于该第一向量加密后的第一加密向量，以及该第二向量加密后的第二加密向量，基于第一加密向量和第二加密向量计算第一加密距离，并基于第一加密距离以及第一目标私钥，确定第一向量和第二向量的目标距离，基于该目标距离以及预设的第一距离阈值确定第一数据和第二数据是否匹配，即可以在第一数据和第二数据不完全相同时，也能实现第一数据和第二数据的模糊匹配，拓宽了使用场景，且在进行模糊匹配过程中引入了第一目标公钥和第一目标私钥分别进行半同态加密和解密，实现了安全求交，保证了匹配过程的安全性，且整个匹配的过程中，第一数据以及第二数据均未以原始数据的形式离开过对应的第一设备以及第二设备，保证了第一数据和第二数据的安全性，实现了原始数据不出库也能实现模糊匹配，进一步保证了匹配过程的安全性。

实施例 2:

为了确定第一数据对应的第一向量，在上述实施例的基础上，在本申请实施例中，所述将待匹配的第一数据输入到预先训练完成的向量转化模型中，获得所述第一数据对应的第一向量包括:

确定所述第一数据对应的第一目标数据类型;

根据所述第一目标数据类型以及预先保存的数据类型和预先训练完成的向量转化模型的对应关系，确定所述第一数据对应的预先训练完成的第一目标向量转化模型；

将所述第一数据输入到所述预先训练完成的第一目标向量转化模型中，获得所述第一数据对应的所述第一向量。

在本申请实施例中，由于待匹配的第一数据可以为文本数据，比如，该第一数据可以为姓名、性别、地址等，也可以为数字数据，比如，该第一数据可以为身份证号、银行卡号以及准考证号等，也可以为图像数据，比如，该第一数据可以为用于进行人脸识别的图像等，因此，为了针对不用数据类型的第一数据，获得对应的第一向量的预先训练完成的向量转化模型也可以不相同。

具体的，可以在第一设备中保存有数据类型和预先训练完成的向量转化模型的对应关系，根据获取到的待匹配的第一数据对应的第一目标数据类型，采用对应的预先训练完成的向量转化模型，获取该第一数据对应的第一向量；其中，该对应的预先训练完成的向量转化模型也就是预先训练完成的第一目标向量转化模型。

其中，本申请对第一数据对应的第一目标数据类型不作具体限定，示例性的，第一目标数据类型可以为文本类型、或数字类型、或图像类型等。另外，本申请对数据类型对应的向量转化模型也不作具体限定，可以根据需求灵活设置。

示例性的，为了准确的确定将第一数据转换为第一向量的向量转化模型，在上述各实施例的基础上，在本申请实施例中，若所述第一目标数据类型为文本类型，则所述第一目标数据类型对应的预先训练完成的第一目标向量转化模型可以为词向量模型或句向量模型；若所述第一目标数据类型为数字类型，则所述第一目标数据类型对应的预先训练完成的第一目标向量转化模型可以为独热（One-Hot）编码模型；若所述第一目标数据类型为图像类型，所述第一目标数据类型对应的预先训练完成的第一目标向量转化模型可以为图

像向量模型。

具体的，若第一数据为文本数据，即第一数据的第一目标数据类型为文本类型，则根据预先保存的数据类型和预先训练完成的向量转化模型的对应关系，可以确定该第一目标数据类型对应的预先训练完成的第一目标向量转化模型，该第一目标向量转化模型可以为词向量模型或者句向量模型，并可以基于该预先训练完成的词向量模型或者句向量模型，获取该第一数据对应的第一向量。

在一种可能的实施方式中，若第一数据为数字数据，即第一数据的第一目标数据类型为数字类型，则根据预先保存的数据类型和预先训练完成的向量转化模型的对应关系，可以确定该第一目标数据类型对应的预先训练完成的第一目标向量转化模型，该第一目标向量转化模型可以为预先训练完成的One-Hot编码模型，并可以基于该预先训练完成的One-Hot编码模型，获取该第一数据对应的第一向量。

在一种可能的实施方式中，若第一数据为图像数据，即第一数据的第一目标数据类型为图像类型，则根据预先保存的数据类型和预先训练完成的向量转化模型的对应关系，可以确定该第一目标数据类型对应的预先训练完成的第一目标向量转化模型，该第一目标向量转化模型可以为预先训练完成的图像向量模型（图像向量嵌入模型），并可以基于该预先训练完成的图像向量模型，获取该第一数据对应的第一向量。

为方便理解，以预先训练完成的向量转化模型输出的向量的维度为5，预先训练完成的向量转化模型为词向量模型为例进行说明。若第一数据为文本数据，且该第一数据为“某海市某东新区晴某小卖部”则将“某海市某东新区晴某小卖部”输入到预先训练完成的词向量模型中，输出的该“某海市某东新区晴某小卖部”对应的第一向量可以为（1.0，2.0，1.5，2.0，3.5）。

若预先训练完成的向量转化模型为One-Hot编码模型，可以预先针对每个数字设置该对应的独热编码，比如，数字包含0-9，则该0-9各个数字中，0对应的独热编码为0000000001、1对应的独热编码为0000000010、2对应的

独热编码为 0000000100、3 对应的独热编码为 0000001000、4 对应的独热编码为 0000010000、5 对应的独热编码为 0000100000、6 对应的独热编码为 0001000000、7 对应的独热编码为 0010000000、8 对应的独热编码为 0100000000、9 对应的独热编码为 1000000000。将数字数据输入到 One-Hot 编码模型中，该 One-Hot 编码模型输出的第一向量中每个第一分量为该第一数据中对应的每个数字的独热编码。

示例性的，若第一数据为数字数据，该数字数据为“12345”，则将“12345”输入到预先训练完成的词向量模型中，输出的该“12345”对应的第一向量为（0000000010，0000000100，0000001000，0000010000，0000100000）。

其中，在对向量转化模型进行训练时，可以预先标注每个数据和数据对应的标注向量，将每个数据和对应的标注向量输入到原始向量转化模型中，根据原始向量转化模型输出的预测向量及对应的标注向量，对该原始向量转化模型的参数进行调整，当满足收敛条件时，确定向量转化模型训练完成。

在本申请实施例中，无论第一数据和第二数据是数字数据还是文本数据，都能实现模糊匹配，进一步扩宽了应用场景。

实施例 3:

为了使得第二设备也能确定第一数据以及第二数据是否匹配，在上述各实施例的基础上，在本申请实施例中，所述确定所述第一向量和所述第二向量的目标距离之后，所述方法还包括:

将所述目标距离发送给所述第二设备，使所述第二设备基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配。

在一种可能的实施方式中，为了使得第二设备也能确定第一数据与第二数据是否匹配，第一设备可以将确定的第一向量和第二向量的目标距离发送给第二设备，第二设备在接收到第一设备发送的第一向量和第二向量的目标距离后，可以基于该目标距离以及预设的第一距离阈值，确定第一数据与第二数据是否匹配。其中第二设备确定第一数据和第二数据是否匹配的过程与上述实施例中第一设备确定第一数据和第二数据是否匹配的过程可以相同，

在此不再赘述。

在一种可能的实施方式中，第一设备确定了第一数据与第二数据是否匹配，得到第一数据与第二数据是否匹配的匹配结果之后，可以将该匹配结果发送给第二设备。可选的，第二设备可以将第一设备确定的匹配结果与自身确定的匹配结果进行比对，以进一步提高准确性。另外，在一种可能的实施方式中，第二设备也可以不进行确定第一数据与第二数据是否匹配，得到匹配结果的过程，直接利用第一设备确定的匹配结果，从而可以节省能耗。

为方便理解，下面通过一个具体实施例对本申请提供的数据匹配过程进行解释说明：

第一设备可以将待匹配的第一数据输入到部署在第一设备中的预先训练完成的向量转化模型中，获得该第一数据对应的第一向量。同样的，第二设备可以将待匹配的第二数据输入到部署在第二设备中的预先训练完成的向量转化模型中，获得该第二数据对应的第二向量。假设第一数据 U1 对应的第一向量为$(x_1, x_2, x_3\ldots\ldots, x_m)$，第二数据 U5 对应的第二向量为$(y_1, y, y_3\ldots\ldots, y_m)$。

第一设备生成第一目标公私钥对 A (pka1，ska1)，其中，该 pka1 为第一目标公钥，ska1 为第一目标私钥，基于该第一目标公钥对第一向量进行半同态加密生成第一加密向量，第一向量（x_1，x_2，$x_3\ldots\ldots$，x_m）对应的第一加密向量为（$E_{pka1}(x_1)$，$E_{pka1}(x_2)$，$E_{pka1}(x_3)\ldots\ldots$，$E_{pka1}(x_m)$），并将该第一加密向量以及该第一目标公钥发送给第二设备。

第二设备接收到第一设备发送的第一目标公钥以及第一加密向量后，基于该第一目标公钥对第二向量进行半同态加密，获得第二加密向量。具体的，该第二向量$(y_1, y, y_3\ldots\ldots, y_m)$对应的第二加密向量为$(E_{pka1}(y_1), E_{pka1}(y_2), E_{pka1}(y_3)\ldots\ldots, E_{pka1}(y_m))$，第二设备基于该第二加密向量以及接收到的第一加密向量，确定第一加密距离$\sum_{i=1}^{m}(E_{pka1}(x_i^2 - 2x_iy_i + y_i^2))$，并将该第一加密距离并发送给第一设备。

第一设备接收到该第一加密距离$\sum_{i=1}^{m}(E_{pka1}(x_i^2-2x_iy_i+y_i^2))$后，根据自身生成的第一目标公钥对应的第一目标私钥对该第一加密距离进行解密，以确定第一向量和第二向量的目标距离$d=\sqrt{\sum_{i=1}^{m}(x_i^2-2x_iy_i+y_i^2)}$。第一设备可以基于该目标距离及预设的第一距离阈值，确定第一数据与第二数据是否匹配，并且第一设备可以将该第一向量和第二向量的目标距离发送给第二设备，第二设备可以根据第二向量与第一向量的目标距离及预设的第一距离阈值，确定第一数据与第二数据是否匹配。

实施例 4:

为了使得第二设备也能确定第一向量和第二向量的目标距离，在上述各实施例的基础上，所述方法还包括:

接收所述第二设备发送的第三加密向量以及所述第二设备生成的第二目标公钥；其中，所述第三加密向量为所述第二设备采用所述第二目标公钥对所述第二向量进行半同态加密后得到的;

基于所述第二目标公钥对所述第一向量进行半同态加密生成第四加密向量;

基于所述第三加密向量以及所述第四加密向量计算得到第二加密距离，并将所述第二加密距离发送给所述第二设备，以使所述第二设备基于所述第二加密距离以及与所述第二目标公钥对应的第二目标私钥，确定所述第一向量和第二向量的目标距离，并使所述第二设备基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配。

在一种可能的实施方式中，为了使得第二设备也能确定第一向量和第二向量的目标距离，而无需依赖第一设备将其确定的目标距离发送给第二设备，进一步提高安全性，第二设备可以生成第二目标公私钥对，其中，该第二目标公私钥对中包含第二目标公钥和第二目标私钥（为方便描述，将第二设备生成的公钥称为第二目标公钥，将第二设备生成的私钥称为第二目标私钥）。可选的，第二目标公私钥对可以为半同态加密公私钥对，即第二目标公钥可

以为半同态加密公钥，第二目标私钥可以为半同态加密私钥。第二设备可以根据自身生成的第二目标公钥对第二向量进行半同态加密生成第三加密向量。其中，该第二目标公私钥对可以为对称公私钥对，也可以为非对称公私钥对，可以根据需求设置目标公私钥对。其中，生成第二目标公私钥对的过程为现有技术，在此不做赘述。

第二设备可以将第三加密向量以及第二设备生成的第二目标公钥发送给第一设备。第一设备接收到第二设备发送的第三加密向量以及第二目标公钥后，可以基于第二目标公钥对第一向量进行半同态加密生成第四加密向量。

第一设备也可以基于第三加密向量和第四加密向量计算得到第二加密距离，其中，计算第二加密距离与计算第一加密距离的过程类似。例如:

在一种可能的实施方式中，第一设备可以针对第三加密向量中的每个第三加密分量、第四向量中的每个第四分量以及第四加密向量中的每个第四加密分量，通过半同态加密算法，确定该第四加密分量的第四加密平方分量、加密后的该第四分量与相应的第三加密分量的乘积，以及相应的第三加密分量的第三加密平方分量；基于该第四加密平方分量、该乘积以及该第三加密平方分量，确定该第四加密分量对应的加密后的子距离；并可以基于每个第四加密分量对应的加密后的子距离，确定第三加密向量和第四加密向量的第二加密距离。为方便理解，下面以公式形式对本申请实施例提供的确定第二加密距离的过程进行解释说明。

示例性的，假设第一数据 U1 对应的第一向量为（x_1，x_2，x_3……，x_m），第二数据 U2 对应的第二向量为（y_1，y，y_3……，y_m）。

第二设备生成第二目标公私钥对 B(pka2，ska2)，其中，该 pka2 为第二目标公钥，ska2 为第二目标私钥，基于该第二目标公钥对第二向量进行半同态加密生成第三加密向量，例如第二向量（y_1，y，y_3……，y_m）对应的第三加密向量为（$E_{pka2}(y_1)$，$E_{pka2}(y_2)$，$E_{pka2}(y_3)$……，$E_{pka2}(y_m)$），第三加密向量中包含 m 个第三加密分量$E_{pka2}(x_i)$，其中，i 为不大于 m 的任一正整数。

第二设备可以将该第三加密向量以及该第二目标公钥发送给第一设备。

第一设备接收到第二设备发送的第二目标公钥以及第三加密向量后，可以基于该第二目标公钥对第一向量进行半同态加密，获得第四加密向量。例如第一向量($x_1, x_2, x_3 \ldots\ldots, x_m$)，对应的第四加密向量为($E_{pka2}(x_1), E_{pka2}(x_2), E_{pka2}(x_3) \ldots\ldots, E_{pka2}(x_m)$)，第四加密向量中包含 m 个第四加密分量 $E_{pka2}(x_i)$，其中，i 为不大于 m 的任一正整数。

在一种可能的实施方式中，针对第四加密向量中的每个第四加密分量 $E_{pka2}(x_i)$，第一设备可以确定该第四加密分量的第四加密平方分量: $E_{pka2}(x_i^2)$。另外，针对第四加密向量中的每个第四加密分量 $E_{pka2}(x_i)$，第一设备还可以确定加密后的第四分量与相应的第三加密分量的乘积，以及相应的第三加密分量的第三加密平方分量: $E_{pka2}(y_i^2)$。

其中，可以采用上述实施例中提供的通过半同态加密算法，确定第一加密分量与相应的第二分量的乘积的方式，来确定第四分量与相应的第三加密分量的乘积。例如，可以先通过半同态加密算法，确定加密后的指数幂 $E_{pka2}(y_i)^{x_i}$，其中，该指数幂的底数可以为该第三加密分量，该指数幂的指数为相应的第四分量 x_i; 该指数幂用于确定基于半同态加密算法加密后的第四分量与相应的第三加密分量的乘积，在此不再赘述。

在一种可能的实施方式中，针对每个第四加密分量，可以基于该第四加密分量对应的第四加密平方分量 $E_{pka2}(x_i^2)$、该第四分量与相应的第三加密分量的乘积 $E_{pka2}(y_i)^{x_i}$ 以及第三加密平方分量 $E_{pka2}(y_i^2)$，确定该第四加密分量对应的加密后的子距离: $E_{pka2}(x_i^2)$-$2E_{pka2}(y_i)^{x_i}$+$E_{pka2}(y_i^2)$。可选的，可以将每个第四加密分量对应的加密后的子距离的和值，确定为第三加密向量与第四加密向量的第二加密距离: $\sum_{i=1}^{m}(E_{pka2}(x_i^2) - 2E_{pka2}(y_i)^{x_i} + E_{pka2}(y_i^2))$。

在一种可能的实施方式中，第一设备获得第二加密距离之后，可以将该第二加密距离发送给第二设备。第二设备接收到该第二加密距离之后，可以根据第二设备自身生成的第二目标私钥对中的第二目标私钥，对该第二加密

距离进行解密，以确定第一向量和第二向量的目标距离。

示例性的，第二设备可以根据第二目标私钥，对第二加密距离进行解密后，得到$\sum_{i=1}^{m}(x_i^2-2x_iy_i+y_i^2)$，可以基于解密后的第二加密距离以及欧氏距离公式，计算第一向量和第二向量之间的目标距离，示例性的，第一向量和第二向量之间的目标距离$d=\sqrt{\sum_{i=1}^{m}(x_i^2-2x_iy_i+y_i^2)}$。其中，该$x_i$为第一向量中的第 i 个分量，$y_i$为第二向量中的第 i 个分量，m 为第一向量或第二向量中包含的分量的数量，且该第一向量中包含的分量的数量与该第二分量中包含的分量的数量相同，即第一向量的长度与第二向量的长度相等，第一向量和第二向量的长度可以均为预设长度 m，m 可以为任意正整数。

另外，本申请实施例还提供了另一种确定第二加密距离的方式，在一种可能的实施方式中，第二加密距离还可以采用以下过程获得:

针对第四加密向量中的每个第四加密分量$E_{pka2}(x_i)$以及第三加密向量中的每个第三加密分量$E_{pka2}(y_i)$，第一设备可以确定该第四分量与相应的第三加密分量的乘积$E_{pka2}(y_i)^{x_i}$，以及相应的第四加密分量的第四加密平方分量$E_{pka2}(x_i^2)$；第一设备可以基于该乘积$E_{pka2}(y_i)^{x_i}$以及该第四加密平方分量$E_{pka2}(x_i^2)$，确定该第四加密分量对应的加密后的子距离：$a=E_{pka2}(x_i)^2-2E_{pka2}(y_i)^{x_i}$。

第一设备可以将该加密后的子距离 a 发送给第二设备，第二设备针对第三加密向量中的每个第三加密分量$E_{pka2}(y_i)$，可以确定该第三加密分量的第三加密平方分量$E_{pka2}(y_i^2)$，并可以基于该第三加密平方分量$E_{pka2}(y_i^2)$以及加密后的子距离$E_{pka2}(x_i)^2-2E_{pka2}(y_i)^{x_i}$，对该第四加密分量对应的加密后的子距离进行更新，可选的，可以将加密后的子距离更新为:

$E_{pka2}(x_i^2)$-$2E_{pka2}(y_i)^{x_i}$+$E_{pka2}(y_i^2)$。

第二设备可以基于每个第四加密分量对应的更新后的子距离，确定第三加密向量和第四加密向量的更新后的第二加密距离。可选的，可以将每个第四加密分量对应的更新后的子距离的和值，确定为第三加密向量和第四加密

向量的更新后的第二加密距离：$\sum_{i=1}^{m}(E_{pka2}(x_i^2 - 2x_iy_i + y_i^2))$。

与上述实施例类似，确定了更新后的第二加密距离之后，第二设备可以基于第二目标私钥，对该更新后的第二加密距离进行解密，以确定第一向量和第二向量的目标距离。示例性的，第二设备可以根据第二目标私钥，对第二加密距离进行解密后，得到$\sum_{i=1}^{m}(x_i^2 - 2x_iy_i + y_i^2)$，可以基于解密后的第二加密距离以及欧氏距离公式，计算第一向量和第二向量之间的目标距离，示例性的，第一向量和第二向量之间的目标距离$d = \sqrt{\sum_{i=1}^{m}(x_i^2 - 2x_iy_i + y_i^2)}$。其中，该$x_i$为第一向量中的第 i 个分量，$y_i$为第二向量中的第 i 个分量，m 为第一向量或第二向量中包含的分量的数量，且该第一向量中包含的分量的数量与该第二分量中包含的分量的数量相同，即第一向量的长度与第二向量的长度相等，第一向量和第二向量的长度可以均为预设长度 m，m 可以为任意正整数。

由于第一数据对应的第一向量的长度与第二数据对应的第二向量的长度相等，因此，即使第一数据与第二数据不相同，也能实现模糊匹配，拓宽了使用场景。

第二设备确定了第一向量和第二向量的目标距离之后，也可以基于该目标距离以及预设的第一距离阈值，确定第一数据和第二数据是否匹配，其中第二设备确定第一数据和第二数据是否匹配的过程与上述实施例中第一设备确定第一数据和第二数据是否匹配的过程可以相同，在此不再赘述。

为方便理解，下面通过一个具体实施例对本申请提供的数据匹配过程进行说明。图 2 为本申请实施例提供的第二种数据匹配过程示意图，如图 2 所示，该过程包括以下操作：

第一设备分别将待匹配的每个第一数据输入到部署在第一设备中的预先训练完成的向量转化模型中，针对每个第一数据，获得该第一数据对应的第一向量。同样的，第二设备分别将待匹配的每个第二数据输入到部署在第二设备中的预先训练完成的向量转化模型中，针对每个第二数据，获得该第二

数据对应的第二向量。如图 2 所述，分别有 4 个第一数据，分别为 U1、U2、U3、U4，且 U1 对应的第一向量为（x11，x12，x13……，x1m），U2 对应的第一向量为（x21，x22，x23……，x2m），U3 对应的第一向量为（x31，x32，x33……，x3m），U4 对应的第一向量为（x41，x42，x43……，x4m）。另外，分别有 4 个第二数据，分别为 U5、U6、U7、U8，且 U5 对应的第二向量为（y11，y12，y13……，y1m），U6 对应的第二向量为（y21，y22，y23……，y2m），U7 对应的第二向量为（y31，y32，y33……，y3m），U8 对应的第二向量为（y41，y42，y43……，y4m）。

第一设备生成第一目标公私钥对 A(pka，ska)，其中，该 pka 为第一目标公钥，ska 为第一目标私钥，该第一目标公私钥对为半同态加密目标公私钥对。针对每个第一向量，基于该第一目标公钥对对该第一向量进行半同态加密，生成相应的第一加密向量。例如，第一向量（x11，x12，x13……，x1m）对应的第一加密向量为（E_{pka}(x11)，E_{pka}(x12)，E_{pka}(x13)……，E_{pka}(x1m)）。第一向量（x21，x22，x23……，x2m）对应的第一加密向量为（E_{pka}(x21)，E_{pka}(x22)，E_{pka}(x23)……，E_{pka}(x2m)）。第一向量（x31，x32，x33……，x3m）对应的第一加密向量为（E_{pka}(x31)，E_{pka}(x32)，E_{pka}(x33)……，E_{pka}(x3m)）。第一向量（x41，x42，x43……，x4m）对应的第一加密向量为（E_{pka}(x41)，E_{pka}(x42)，E_{pka}(x43)……，E_{pka}(x4m)）。

第一设备将第一目标公钥以及每个第一加密向量均发送给第二设备，第二设备在接收到该第一目标公钥以及每个第一加密向量后，针对每个第二向量，基于第一目标公钥对该第二向量进行半同态加密生成第二加密向量。第二向量（y11，y12，y13……，y1m）对应的第二加密向量为（E_{pka}(y11)，E_{pka}(y12)，E_{pka}(y13)……，E_{pka}(y1m)）。第二向量（y21，y22，y23……，y2m）对应的第二加密向量为（E_{pka}(y21)，E_{pka}(y22)，E_{pka}(y23)……，E_{pka}(y2m)）。第二向量（y31，y32，y33……，y3m）对应的第二加密向量为（E_{pka}(y31)，E_{pka}(y32)，E_{pka}(y33)……，E_{pka}(y3m)）。第

二向量（y41，y42，y43……，y4m）对应的第二加密向量为（E_{pka}(y41)，E_{pka}(y42)，E_{pka}(y43)……，E_{pka}(y4m)）。

由于针对每个第一数据，需要确定该第一数据与每个第二数据是否匹配，例如，针对第一数据 U1，需要确定第一数据 U1 与第二数据 U5 是否匹配、第一数据U1与第二数据U6是否匹配、第一数据U1与第二数据U7是否匹配、第一数据 U1 与第二数据 U8 是否匹配，因此，针对每个第一向量，均可以确定该第一向量与每个第二向量的目标距离，针对每个第一加密向量，也均可以基于该第一加密向量与每个第二加密向量，分别计算相应的第一加密距离。其中，基于任一第一加密向量与任一第二加密向量，计算第一加密距离的过程均与上述实施例中计算第一加密距离的过程相同，例如，针对任一第一加密向量与任一第二加密向量，第二设备可以将基于该第一加密向量和该第二加密向量计算的加密后的子距离 b，$(-2E_{pka1}(x_i)^{y_i} + E_{pka1}(y_i^2))$ 发送给第一设备。第一设备可以将加密后的子距离更新为 $E_{pka1}(x_i^2)$-$2E_{pka1}(x_iy_i)$+$E_{pka1}(y_i^2)$，并采用第一目标私钥，对更新后的第一加密距离进行解密，以进而得到相应的第一向量与第二向量的目标距离，在此不再赘述。

针对每个第一向量，分别确定了该第一向量与每个第二向量的目标距离之后，第一设备还可以将确定的目标距离发送给第二设备，第一设备和第二设备均可以分别基于该目标距离以及预设的第一距离阈值，确定该第一向量对应的第一数据分别与每个第二向量对应的第二数据是否匹配，即可以分别确定各个第一数据与各个第二数据是否匹配。

为方便理解，下面再通过一个具体实施例对本申请提供的数据匹配过程进行说明。假设保存在第一设备中的第一数据有 3 个，分别为“某海市某东新区晴某小卖部”、“某海市天某餐饮店”、“高某路杨某福麻辣烫”。保存在第二设备中的第二数据也有 3 个，分别为“某海市晴某小卖部”、“某海市天某餐饮店”、“麦某劳”。

其中第一数据中，“某海市某东新区晴某小卖部”对应的第一向量为＜1.0,2.0,1.5,2.0,3.5＞，记为 A1。“某海市天某餐饮店”对应的第一向量为＜3.0,4.0,2.5,2.5,1.5＞，记为 A2。“高某路杨某福麻辣烫”对应的第一向量为＜4.5,5.5,7.5,1.5,0.5＞，记为 A3。

第二数据中，“某海市晴某小卖部”对应的第二向量为＜1.0,2.0,1.5,1.0,3.5＞，记为 B1。“某海市天某餐饮店”对应的第二向量为＜3.0,4.0,2.5,2.5,1.5＞，记为 B2。“麦某劳”对应的第二向量为＜3.5,6.5,2.5,7.5,2.5＞，记为 B3。

第一设备生成第一目标公私钥对 A(pka，ska)，其中，该 pka 为第一目标公钥，ska 为第一目标私钥，该第一目标公私钥对为半同态加密目标公私钥对。针对每个第一向量，第一设备可以基于该第一目标公钥对对该第一向量进行半同态加密，生成相应的第一加密向量。

第一设备将第一目标公钥以及每个第一加密向量均发送给第二设备，第二设备在接收到该第一目标公钥以及每个第一加密向量后，针对每个第二向量，基于第一目标公钥对该第二向量进行半同态加密生成第二加密向量。

由于针对每个第一数据，需要确定该第一数据与每个第二数据是否匹配，因此，针对每个第一向量，均可以确定该第一向量与每个第二向量的目标距离，针对每个第一加密向量，也均可以基于该第一加密向量与每个第二加密向量，分别计算相应的第一加密距离。其中，基于任一第一加密向量与任一第二加密向量，计算第一加密距离的过程均与上述实施例中计算第一加密距离的过程相同，在此不再赘述。

如表 1 所示，表 1 为本申请一些实施例提供的一种目标距离的示意表。

表 1

D(x,y)	B1	B2	B3
A1	1	3.64	7.66
A2	3.9	0	5.7

A3	8.35	5.61	8.18

假设 D(x,y)表征目标距离，其中 x 表征的第一向量，y 表征的第二向量。A1 与 B1 对应的目标距离为 1，A1 与 B2 对应的目标距离为 3.64，A1 与 B3 对应的目标距离为 7.66，A2 与 B1 对应的目标距离为 3.9，A2 与 B2 对应的目标距离为 0，A2 与 B3 对应的目标距离为 5.7，A3 与 B1 对应的目标距离为 8.35，A3 与 B2 对应的目标距离为 5.16，A3 与 B3 对应的目标距离为 8.18。

假设第一距离阈值为 2，则本次模糊匹配结果为第一数据中的“某海市某东新区晴某小卖部”与第二数据中的“某海市晴某小卖部”匹配，第一数据中的“某海市天某餐饮店”与第二数据中的“某海市天某餐饮店”匹配，由此也可以看出，本申请中的数据匹配方法可以使得两个数据在不完全相同时，也能实现数据的模糊匹配，拓宽了使用场景。

为方便理解，下面再通过一个具体实施例对本申请提供的数据匹配过程进行说明。假设保存在第一设备中的第一数据有 3 个，分别为 3 个手机号码“13345678909”、“13245678911”、“13334536787”。保存在第二设备中的第二数据也有 3 个，分别为 3 个手机号码“13334536787”、“13345678908”、“15439402290”。

第一设备和第二设备分别可以利用独热（One-Hot）编码模型（数据转化模型）第一数据和第二数据中的手机号码进行向量化，生成相应的第一向量和第二向量，可选的，第一向量和第二向量的向量维度可以为 10*11 维，其中 10 代表 0-9 共 10 个数字，11 代表手机号长度为 11 个数字。示例性的，参阅表 2，表 2 为手机号码“13345678909”对应的向量示意表。

表 2

手机号码中的其中一个数字	数字对应的向量（独热编码）
1	0000000010
3	0000001000
3	0000001000
4	0000010000
5	0000100000
6	0001000000

7	0010000000
8	0100000000
9	1000000000
0	0000000001
9	1000000000

参阅表 2，在一种可能的实施方式中，0 对应的独热编码为 0000000001、1 对应的独热编码为 0000000010、2 对应的独热编码为 0000000100、3 对应的独热编码为 0000001000、4 对应的独热编码为 0000010000、5 对应的独热编码为 0000100000、6 对应的独热编码为 0001000000、7 对应的独热编码为 0010000000、8 对应的独热编码为 0100000000、9 对应的独热编码为 1000000000。则手机号码“13345678909”对应的第一向量（或第二向量）可以为（0000000010，0000001000，0000001000，0000010000，0000100000，0001000000，0010000000，0100000000，1000000000，0000000001，1000000000）。确定其他手机号码对应的向量的过程与该过程类似，在此不再赘述。

为方便描述，将第一数据中“13345678909”对应的第一向量记为 A4、“13245678911”对应的第一向量记为 A5、“13334536787”对应的第一向量记为 A6。将第二数据中“13334536787”对应的第二向量记为 B4、“13345678908”对应的第二向量记为 B5、“15439402290”对应的第二向量记为 B6。

第一设备生成第一目标公私钥对 A(pka，ska)，其中，该 pka 为第一目标公钥，ska 为第一目标私钥，该第一目标公私钥对为半同态加密目标公私钥对。针对每个第一向量，第一设备可以基于该第一目标公钥对对该第一向量进行半同态加密，生成相应的第一加密向量。

第一设备将第一目标公钥以及每个第一加密向量均发送给第二设备，第二设备在接收到该第一目标公钥以及每个第一加密向量后，针对每个第二向量，基于第一目标公钥对该第二向量进行半同态加密生成第二加密向量。

由于针对每个第一数据，需要确定该第一数据与每个第二数据是否匹配，因此，针对每个第一向量，均可以确定该第一向量与每个第二向量的目标距离，针对每个第一加密向量，也均可以基于该第一加密向量与每个第二加密向量，分别计算相应的第一加密距离。其中，基于任一第一加密向量与任一

第二加密向量，计算第一加密距离的过程均与上述实施例中计算第一加密距离的过程相同，在此不再赘述。

如表 3 所示，表 3 为本申请一些实施例提供的一种目标距离的示意表。

表 3

D(x,y)	B4	B5	B6
A4	2.82	1	3.16
A5	3	1.73	3
A6	0	2.82	3

假设 D(x,y)表征目标距离，其中 x 表征的第一向量，y 表征的第二向量。A4 与 B4 对应的目标距离为 2.82，A4 与 B5 对应的目标距离为 1，A4 与 B6 对应的目标距离为 3.16，A5 与 B4 对应的目标距离为 3，A5 与 B5 对应的目标距离为 1.73，A5 与 B6 对应的目标距离为 3，A6 与 B4 对应的目标距离为 0，A6 与 B5 对应的目标距离为 2.83，A6 与 B6 对应的目标距离为 3。

假设第一距离阈值为 2，则本次模糊匹配结果为第一数据中的“13345678909”与第二数据中的“13345678909”匹配，第一数据中的“13334536787”与第二数据中的“13334536787”匹配，由此也可以看出，本申请中的数据匹配方法不仅能实现数据的模糊匹配，还可以实现数据的精准匹配，拓宽了使用场景。

实施例 5:

基于相同的技术构思，本申请提供了一种数据匹配方法，应用于第二设备，图 3 为本申请实施例提供的第三种数据匹配过程示意图，如图 3 所示，该方法包括:

S301：将待匹配的第二数据输入到预先训练完成的向量转化模型中，获得所述第二数据对应的第二向量。

本申请实施例提供的数据匹配方法应用于第二设备，该第二设备可以是智能终端、PC 或者服务器等设备，且该第二设备与本申请中的第一设备为不同的设备。

S302：接收第一设备发送的第一目标公钥以及第一加密向量，采用所述第一目标公钥对所述第二向量进行半同态加密生成第二加密向量；其中，所述第一加密向量为采用所述第一目标公钥对第一向量进行半同态加密后得到的，所述第一向量为将第一数据输入到所述第一设备中的预先训练完成的向量转化模型中获得的。

S303：基于所述第一加密向量和第二加密向量计算第一加密距离，将所述第一加密距离发送给所述第一设备，以使所述第一设备基于所述第一加密距离以及与所述第一目标公钥对应的第一目标私钥，确定所述第一向量和第二向量的目标距离；并使所述第一设备基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配。

在一种可能的实施方式中，所述将待匹配的第二数据输入到预先训练完成的向量转化模型中，获得所述第二数据对应的第二向量包括：

确定所述第二数据对应的第二目标数据类型；

根据所述第二目标数据类型以及预先保存的数据类型和预先训练完成的向量转化模型的对应关系，确定所述第二数据对应的预先训练完成的第二目标向量转化模型；

将所述第二数据输入到所述预先训练完成的第二目标向量转化模型中，获得所述第二数据对应的所述第二向量。

在一种可能的实施方式中，所述第二目标数据类型为文本类型、数字类型、图像类型中的至少一种。

在一种可能的实施方式中，若所述第二目标数据类型为文本类型，所述第二目标数据类型对应的预先训练完成的第二目标向量转化模型为词向量模型或句向量模型；若所述第二目标数据类型为数字类型，所述第二目标数据类型对应的预先训练完成的第二目标向量转化模型为独热编码模型；若所述第一目标数据类型为图像类型，所述第一目标数据类型对应的预先训练完成的第一目标向量转化模型为图像向量模型。

在一种可能的实施方式中，所述方法还包括：

采用自身生成的第二目标公钥对所述第二向量进行半同态加密生成第三加密向量，将所述第二目标公钥和所述第三加密向量发送给所述第一设备；

接收所述第一设备发送的基于所述第三加密向量以及第四加密向量计算得到的第二加密距离，所述第四加密向量为采用所述第二目标公钥对第一向量进行半同态加密后得到的；所述第一向量为将第一数据输入到所述第一设备中的预先训练完成的向量转化模型中获得的；

基于所述第二加密距离以及与所述第二目标公钥对应的第二目标私钥，确定所述第一向量和第二向量的目标距离，并使所述第二设备基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配。

在一种可能的实施方式中，所述将所述第一加密距离发送给所述第一设备之后，所述方法还包括：

接收所述第一设备发送的目标距离，基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配。

在一种可能的实施方式中，所述方法还包括：

接收所述第一设备发送的所述第一数据以及所述第二数据是否匹配的匹配结果。

在一种可能的实施方式中，所述采用所述第一目标公钥对所述第二向量进行半同态加密生成第二加密向量包括：

针对所述第二向量中的每个第二分量，基于所述第一目标公钥对所述第二向量中的每个第二分量分别进行半同态加密，生成所述第二加密向量。

实施例 6：

基于相同的技术构思，本申请提供了一种数据匹配装置，应用于第一设备。图 4 为本申请一些实施例提供的一种数据匹配装置结构示意图，如图 4 所示，所述装置包括：

第一获取模块 41，用于将待匹配的第一数据输入到预先训练完成的向量转化模型中，获得所述第一数据对应的第一向量；

第一处理模块 42，用于采用自身生成的第一目标公钥对所述第一向量进行半同态加密生成第一加密向量，将所述第一目标公钥以及所述第一加密向量发送给第二设备；

所述第一获取模块 41，还用于获取所述第二设备发送的基于所述第一加密向量和第二加密向量计算的第一加密距离，所述第二加密向量为采用所述第一目标公钥对第二向量进行半同态加密后得到的；所述第二向量为将第二数据输入到所述第二设备中的预先训练完成的向量转化模型中获得的；

第一确定模块 43，用于基于所述第一加密距离以及与所述第一目标公钥对应的第一目标私钥，确定所述第一向量和第二向量的目标距离；基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配。

在一种可能的实施方式中，所述第一获取模块 41，具体用于确定所述第一数据对应的第一目标数据类型；

根据所述第一目标数据类型以及预先保存的数据类型和预先训练完成的向量转化模型的对应关系，确定所述第一数据对应的预先训练完成的第一目标向量转化模型；

将所述第一数据输入到所述预先训练完成的第一目标向量转化模型中，获得所述第一数据对应的所述第一向量。

在一种可能的实施方式中，所述第一目标数据类型为文本类型、数字类型、图像类型中的至少一种。

在一种可能的实施方式中，若所述第一目标数据类型为文本类型，所述第一目标数据类型对应的预先训练完成的第一目标向量转化模型为词向量模型或句向量模型；若所述第一目标数据类型为数字类型，所述第一目标数据类型对应的预先训练完成的第一目标向量转化模型为独热编码模型；若所述第一目标数据类型为图像类型，所述第一目标数据类型对应的预先训练完成的第一目标向量转化模型为图像向量模型。

在一种可能的实施方式中，所述装置还包括：

接收模块，用于接收所述第二设备发送的第三加密向量以及所述第二设备生成的第二目标公钥；其中，所述第三加密向量为所述第二设备采用所述第二目标公钥对所述第二向量进行半同态加密后得到的；

所述第一处理模块，还用于基于所述第二目标公钥对所述第一向量进行半同态加密生成第四加密向量；

发送模块，用于基于所述第三加密向量以及所述第四加密向量计算得到第二加密距离，并将所述第二加密距离发送给所述第二设备，以使所述第二设备基于所述第二加密距离以及与所述第二目标公钥对应的第二目标私钥，确定所述第一向量和第二向量的目标距离，并使所述第二设备基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配。

在一种可能的实施方式中，所述第一发送模块，还用于将所述目标距离发送给所述第二设备，使所述第二设备基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配。

在一种可能的实施方式中，所述第一发送模块，还用于将确定的所述第一数据以及所述第二数据是否匹配的匹配结果发送给所述第二设备。

在一种可能的实施方式中，所述第一处理模块，具体用于针对所述第一向量中的每个第一分量，基于所述第一目标公钥对所述第一向量中的每个第一分量分别进行半同态加密，生成所述第一加密向量。

在一种可能的实施方式中，所述第一向量和所述第二向量的长度均为预设长度。

在一种可能的实施方式中，所述第一加密距离是采用以下过程获得的：

针对所述第一加密向量中的每个第一加密分量以及所述第二加密向量中的每个第二加密分量，通过半同态加密算法，确定该第一加密分量与相应的第二加密分量的乘积，以及相应的第二加密分量的第二加密平方分量；基于该乘积以及该第二加密平方分量，确定该第一加密分量对应的加密后的子距离；

基于每个第一加密分量对应的加密后的子距离，确定所述第一加密向量和第二加密向量的第一加密距离。

在一种可能的实施方式中，所述第一确定模块 43，具体用于针对所述第一加密向量中的每个第一加密分量，确定该第一加密分量的第一加密平方分量；基于该第一加密平方分量以及该第一加密分量对应的所述加密后的子距离，对该第一加密分量对应的所述加密后的子距离进行更新；

基于每个第一加密分量对应的更新后的子距离，确定所述第一加密向量和第二加密向量的更新后的第一加密距离；

采用与所述第一目标公钥对应的第一目标私钥，对所述更新后的第一加密距离进行解密，得到所述第一向量和第二向量的目标距离。

在一种可能的实施方式中，所述第一加密距离是采用以下过程获得的：

针对所述第一加密向量中的每个第一加密分量以及所述第二加密向量中的每个第二加密分量，通过半同态加密算法，确定该第一加密分量的第一加密平方分量、该第一加密分量与相应的第二加密分量的乘积，以及相应的第二加密分量的第二加密平方分量；基于该第一加密平方分量、该乘积以及该第二加密平方分量，确定该第一加密分量对应的加密后的子距离；

基于每个第一加密分量对应的加密后的子距离，确定所述第一加密向量和第二加密向量的第一加密距离。

在一种可能的实施方式中，所述第一确定模块 43，具体用于确定该第一加密分量个指数幂的乘积，其中，所述指数幂的指数为相应的第二加密分量；

基于该第一加密分量个指数幂的乘积，确定该第一加密分量个相应的第二加密分量的和；

基于该第一加密分量个相应的第二加密分量的和，确定该第一加密分量与相应的第二加密分量的乘积。

在一种可能的实施方式中，所述第一确定模块 43，具体用于判断所述目标距离是否小于预设的第一距离阈值；

若是，则确定所述第一数据与所述第二数据匹配；

否则，确定所述第一数据与所述第二数据不匹配。

在一种可能的实施方式中，所述第一确定模块 43，还用于判断所述目标距离是否等于预设的第二距离阈值，若是，则确定所述第一数据与所述第二数据相同。

实施例 7:

基于相同的技术构思，本申请提供了另一种数据匹配装置，应用于第二设备。图 5 为本申请一些实施例提供的另一种数据匹配装置结构示意图，如图 5 所示，所述装置包括:

第二获取模块 51，用于将待匹配的第二数据输入到预先训练完成的向量转化模型中，获得所述第二数据对应的第二向量;

第二处理模块 52，用于接收第一设备发送的第一目标公钥以及第一加密向量，采用所述第一目标公钥对所述第二向量进行半同态加密生成第二加密向量；其中，所述第一加密向量为采用所述第一目标公钥对第一向量进行半同态加密后得到的，所述第一向量为将第一数据输入到所述第一设备中的预先训练完成的向量转化模型中获得的;

第二确定模块 53，用于基于所述第一加密向量和第二加密向量计算第一加密距离，将所述第一加密距离发送给所述第一设备，以使所述第一设备基于所述第一加密距离以及与所述第一目标公钥对应的第一目标私钥，确定所述第一向量和第二向量的目标距离；并使所述第一设备基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配。

在一种可能的实施方式中，所述第二获取模块 51，具体用于确定所述第二数据对应的第二目标数据类型;

根据所述第二目标数据类型以及预先保存的数据类型和预先训练完成的向量转化模型的对应关系，确定所述第二数据对应的预先训练完成的第二目标向量转化模型;

将所述第二数据输入到所述预先训练完成的第二目标向量转化模型中，获得所述第二数据对应的所述第二向量。

在一种可能的实施方式中，所述第二目标数据类型为文本类型、数字类型、图像类型中的至少一种。

在一种可能的实施方式中，若所述第二目标数据类型为文本类型，所述第二目标数据类型对应的预先训练完成的第二目标向量转化模型为词向量模型或句向量模型；若所述第二目标数据类型为数字类型，所述第二目标数据类型对应的预先训练完成的第二目标向量转化模型为独热编码模型；若所述第一目标数据类型为图像类型，所述第一目标数据类型对应的预先训练完成的第一目标向量转化模型为图像向量模型。

在一种可能的实施方式中，所述第二处理模块 52，还用于采用自身生成的第二目标公钥对所述第二向量进行半同态加密生成第三加密向量，将所述第二目标公钥和所述第三加密向量发送给所述第一设备；接收所述第一设备发送的基于所述第三加密向量以及第四加密向量计算得到的第二加密距离，所述第四加密向量为采用所述第二目标公钥对第一向量进行半同态加密后得到的；所述第一向量为将第一数据输入到所述第一设备中的预先训练完成的向量转化模型中获得的；基于所述第二加密距离以及与所述第二目标公钥对应的第二目标私钥，确定所述第一向量和第二向量的目标距离，并使所述第二设备基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配。

在一种可能的实施方式中，所述第二确定模块 53，还用于接收所述第一设备发送的目标距离，基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配。

在一种可能的实施方式中，所述第二确定模块 53，还用于接收所述第一设备发送的所述第一数据以及所述第二数据是否匹配的匹配结果。

在一种可能的实施方式中，所述第二获取模块 51，还用于针对所述第二向量中的每个第二分量，基于所述第一目标公钥对所述第二向量中的每个第二分量分别进行半同态加密，生成所述第二加密向量。

实施例 8:

基于相同的技术构思，本申请提供了一种数据匹配系统，图 6 为本申请一些实施例提供的一种数据匹配系统结构示意图，如图 6 所示，所述系统包括:

第一设备 61，用于将待匹配的第一数据输入到预先训练完成的向量转化模型中，获得所述第一数据对应的第一向量；采用自身生成的第一目标公钥对所述第一向量进行半同态加密生成第一加密向量，将所述第一目标公钥以及所述第一加密向量发送给第二设备 62;

所述第二设备 62，用于将待匹配的第二数据输入到预先训练完成的向量转化模型中，获得所述第二数据对应的第二向量；并接收所述第一设备发送的第一目标公钥以及第一加密向量，采用所述第一目标公钥对所述第二向量进行半同态加密生成第二加密向量；基于所述第一加密向量和第二加密向量计算第一加密距离，将所述第一加密距离发送给所述第一设备 61;

所述第一设备 61，还用于基于所述第一加密距离以及与所述第一目标公钥对应的第一目标私钥，确定所述第一向量和第二向量的目标距离；基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配。

基于相同的技术构思，本申请还提供了一种电子设备，图 7 为本申请一些实施例提供的一种电子设备结构示意图，如图 7 所示，包括：处理器 71、通信接口 72、存储器 73 和通信总线 74，其中，处理器 71，通信接口 72，存储器 73 通过通信总线 74 完成相互间的通信;

所述存储器 73 中存储有计算机程序，当所述程序被所述处理器 71 执行时，使得所述处理器 71 执行上述任一实施例的数据匹配方法的步骤。

由于上述电子设备解决问题的原理与数据匹配方法相似，因此上述电子设备的实施可以参见方法的实施，重复之处不再赘述。

上述电子设备提到的通信总线可以是外设部件互连标准（Peripheral Component Interconnect，PCI）总线或扩展工业标准结构（Extended Industry Standard Architecture，EISA）总线等。该通信总线可以分为地址总线、数据

总线、控制总线等。为便于表示，图中仅用一条粗线表示，但并不表示仅有一根总线或一种类型的总线。

通信接口 72 用于上述电子设备与其他设备之间的通信。

存储器可以包括随机存取存储器（Random Access Memory，RAM），也可以包括非易失性存储器（Non-Volatile Memory，NVM），例如至少一个磁盘存储器。可选地，存储器还可以是至少一个位于远离前述处理器的存储装置。

上述处理器可以是通用处理器，包括中央处理器、网络处理器（Network Processor，NP）等；还可以是数字指令处理器（Digital Signal Processing，DSP）、专用集成电路、现场可编程门陈列或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。

基于相同的技术构思，本申请实施例提供了一种计算机可读存储介质，所述计算机可读存储介质内存储有可由电子设备执行的计算机程序，当所述程序在所述电子设备上运行时，使得所述电子设备执行时实现上述任一实施例的数据匹配方法的步骤。由于上述计算机可读存储介质解决问题的原理与数据匹配方法相似，因此上述计算机可读存储介质的实施可以参见方法的实施，重复之处不再赘述。

上述计算机可读存储介质可以是电子设备中的处理器能够存取的任何可用介质或数据存储设备，包括但不限于磁性存储器如软盘、硬盘、磁带、磁光盘（MO）等、光学存储器如 CD、DVD、BD、HVD 等、以及半导体存储器如 ROM、EPROM、EEPROM、非易失性存储器（NAND FLASH）、固态硬盘（SSD）等。

基于相同的技术构思，在上述各实施例的基础上，本申请提供了一种计算机程序产品，所述计算机程序产品包括：计算机程序代码，当所述计算机程序代码在计算机上运行时，使得计算机执行如上述任一所述数据匹配方法的步骤。

本领域内的技术人员应明白，本申请的实施例可提供为方法、系统、或计算机程序产品。因此，本申请可采用完全硬件实施例、完全软件实施例、

或结合软件和硬件方面的实施例的形式。而且，本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质（包括但不限于磁盘存储器、CD-ROM、光学存储器等）上实施的计算机程序产品的形式。

本申请是参照根据本申请的方法、设备（系统）、和计算机程序产品的流程图和／或方框图来描述的。应理解可由计算机程序指令实现流程图和／或方框图中的每一流程和／或方框、以及流程图和／或方框图中的流程和／或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器，使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和／或方框图一个方框或多个方框中指定的功能的装置。

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

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

显然，本领域的技术人员可以对本申请进行各种改动和变型而不脱离本申请的精神和范围。这样，倘若本申请的这些修改和变型属于本申请权利要求及其等同技术的范围之内，则本申请也意图包含这些改动和变型在内。

权 利 要 求

1、一种数据匹配方法，应用于第一设备，所述方法包括:

将待匹配的第一数据输入到预先训练完成的向量转化模型中，获得所述第一数据对应的第一向量;

采用自身生成的第一目标公钥对所述第一向量进行半同态加密生成第一加密向量，将所述第一目标公钥以及所述第一加密向量发送给第二设备;

获取所述第二设备发送的基于所述第一加密向量和第二加密向量计算的第一加密距离，所述第二加密向量为采用所述第一目标公钥对第二向量进行半同态加密后得到的；所述第二向量为将第二数据输入到所述第二设备中的预先训练完成的向量转化模型中获得的;

基于所述第一加密距离以及与所述第一目标公钥对应的第一目标私钥，确定所述第一向量和第二向量的目标距离;

基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配。

2、根据权利要求 1 所述的方法，其中，所述将待匹配的第一数据输入到预先训练完成的向量转化模型中，获得所述第一数据对应的第一向量包括:

确定所述第一数据对应的第一目标数据类型;

根据所述第一目标数据类型以及预先保存的数据类型和预先训练完成的向量转化模型的对应关系，确定所述第一数据对应的预先训练完成的第一目标向量转化模型;

将所述第一数据输入到所述预先训练完成的第一目标向量转化模型中，获得所述第一数据对应的所述第一向量。

3、根据权利要求 2 所述的方法，其中，所述第一目标数据类型为文本类型、数字类型、图像类型中的至少一种。

4、根据权利要求 3 所述的方法，其中，若所述第一目标数据类型为文本类型，所述第一目标数据类型对应的预先训练完成的第一目标向量转化模型

为词向量模型或句向量模型；若所述第一目标数据类型为数字类型，所述第一目标数据类型对应的预先训练完成的第一目标向量转化模型为独热编码模型；若所述第一目标数据类型为图像类型，所述第一目标数据类型对应的预先训练完成的第一目标向量转化模型为图像向量模型。

5、根据权利要求 1 所述的方法，所述方法还包括:

接收所述第二设备发送的第三加密向量以及所述第二设备生成的第二目标公钥；其中，所述第三加密向量为所述第二设备采用所述第二目标公钥对所述第二向量进行半同态加密后得到的;

基于所述第二目标公钥对所述第一向量进行半同态加密生成第四加密向量;

基于所述第三加密向量以及所述第四加密向量计算得到第二加密距离，并将所述第二加密距离发送给所述第二设备，以使所述第二设备基于所述第二加密距离以及与所述第二目标公钥对应的第二目标私钥，确定所述第一向量和第二向量的目标距离，并使所述第二设备基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配。

6、根据权利要求 1 所述的方法，所述确定所述第一向量和第二向量的目标距离之后，所述方法还包括:

将所述目标距离发送给所述第二设备，使所述第二设备基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配。

7、根据权利要求 1 所述的方法，所述方法还包括:

将确定的所述第一数据以及所述第二数据是否匹配的匹配结果发送给所述第二设备。

8、根据权利要求 1 所述的方法，其中，所述采用自身生成的第一目标公钥对所述第一向量进行半同态加密生成第一加密向量包括:

针对所述第一向量中的每个第一分量，基于所述第一目标公钥对所述第一向量中的每个第一分量分别进行半同态加密，生成所述第一加密向量。

9、根据权利要求 1 或 8 所述的方法，其中，所述第一向量和所述第二向

量的长度均为预设长度。

10、根据权利要求 8 所述的方法，其中，所述第一加密距离是采用以下过程获得的：

针对所述第一加密向量中的每个第一加密分量、所述第二向量中的每个第二分量以及所述第二加密向量中的每个第二加密分量，通过半同态加密算法，确定加密后的该第一加密分量与相应的第二分量的乘积，以及相应的第二加密分量的第二加密平方分量；基于该乘积以及该第二加密平方分量，确定该第一加密分量对应的加密后的子距离；

基于每个第一加密分量对应的加密后的子距离，确定所述第一加密向量和第二加密向量的第一加密距离。

11、根据权利要求 10 所述的方法，其中，所述基于所述第一加密距离以及与所述第一目标公钥对应的第一目标私钥，确定所述第一向量和第二向量的目标距离，包括：

针对所述第一加密向量中的每个第一加密分量，确定该第一加密分量的第一加密平方分量；基于该第一加密平方分量以及该第一加密分量对应的所述加密后的子距离，对该第一加密分量对应的所述加密后的子距离进行更新；

基于每个第一加密分量对应的更新后的子距离，确定所述第一加密向量和第二加密向量的更新后的第一加密距离；

采用与所述第一目标公钥对应的第一目标私钥，对所述更新后的第一加密距离进行解密，得到所述第一向量和第二向量的目标距离。

12、根据权利要求 8 所述的方法，其中，所述第一加密距离是采用以下过程获得的：

针对所述第一加密向量中的每个第一加密分量、所述第二向量中的每个第二分量以及所述第二加密向量中的每个第二加密分量，通过半同态加密算法，确定该第一加密分量的第一加密平方分量、加密后的该第一加密分量与相应的第二分量的乘积，以及相应的第二加密分量的第二加密平方分量；基于该第一加密平方分量、该乘积以及该第二加密平方分量，确定该第一加密

分量对应的加密后的子距离;

基于每个第一加密分量对应的加密后的子距离，确定所述第一加密向量和第二加密向量的第一加密距离。

13、根据权利要求 10 或 12 所述的方法，其中，通过半同态加密算法，确定加密后的该第一加密分量与相应的第二分量的乘积，包括:

通过半同态加密算法，确定加密后的指数幂，其中，所述指数幂的底数为该第一加密分量，所述指数幂的指数为相应的第二分量；所述指数幂用于确定基于半同态加密算法加密后的该第一加密分量与相应的第二分量的乘积。

14、根据权利要求 1 所述的方法，其中，所述基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配包括:

判断所述目标距离是否小于预设的第一距离阈值;

若是，则确定所述第一数据与所述第二数据匹配;

否则，确定所述第一数据与所述第二数据不匹配。

15、根据权利要求 14 所述的方法，所述确定所述第一数据与所述第二数据匹配之后，所述方法还包括:

判断所述目标距离是否等于预设的第二距离阈值，若是，则确定所述第一数据与所述第二数据相同。

16、一种数据匹配方法，应用于第二设备，所述方法包括:

将待匹配的第二数据输入到预先训练完成的向量转化模型中，获得所述第二数据对应的第二向量;

接收第一设备发送的第一目标公钥以及第一加密向量，采用所述第一目标公钥对所述第二向量进行半同态加密生成第二加密向量；其中，所述第一加密向量为采用所述第一目标公钥对第一向量进行半同态加密后得到的，所述第一向量为将第一数据输入到所述第一设备中的预先训练完成的向量转化模型中获得的;

基于所述第一加密向量和第二加密向量计算第一加密距离，将所述第一加密距离发送给所述第一设备，以使所述第一设备基于所述第一加密距离以

及与所述第一目标公钥对应的第一目标私钥，确定所述第一向量和第二向量的目标距离；并使所述第一设备基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配。

17、根据权利要求 16 所述的方法，其中，所述将待匹配的第二数据输入到预先训练完成的向量转化模型中，获得所述第二数据对应的第二向量包括:

确定所述第二数据对应的第二目标数据类型;

根据所述第二目标数据类型以及预先保存的数据类型和预先训练完成的向量转化模型的对应关系，确定所述第二数据对应的预先训练完成的第二目标向量转化模型;

将所述第二数据输入到所述预先训练完成的第二目标向量转化模型中，获得所述第二数据对应的所述第二向量。

18、根据权利要求 17 所述的方法，其中，所述第二目标数据类型为文本类型、数字类型、图像类型中的至少一种。

19、根据权利要求 18 所述的方法，其中，若所述第二目标数据类型为文本类型，所述第二目标数据类型对应的预先训练完成的第二目标向量转化模型为词向量模型或句向量模型；若所述第二目标数据类型为数字类型，所述第二目标数据类型对应的预先训练完成的第二目标向量转化模型为独热编码模型；若所述第一目标数据类型为图像类型，所述第一目标数据类型对应的预先训练完成的第一目标向量转化模型为图像向量模型。

20、根据权利要求 16 所述的方法，所述方法还包括:

采用自身生成的第二目标公钥对所述第二向量进行半同态加密生成第三加密向量，将所述第二目标公钥和所述第三加密向量发送给所述第一设备;

接收所述第一设备发送的基于所述第三加密向量以及第四加密向量计算得到的第二加密距离，所述第四加密向量为采用所述第二目标公钥对第一向量进行半同态加密后得到的；所述第一向量为将第一数据输入到所述第一设备中的预先训练完成的向量转化模型中获得的;

基于所述第二加密距离以及与所述第二目标公钥对应的第二目标私钥，

确定所述第一向量和第二向量的目标距离，并使所述第二设备基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配。

21、根据权利要求 16 所述的方法，所述将所述第一加密距离发送给所述第一设备之后，所述方法还包括:

接收所述第一设备发送的目标距离，基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配。

22、根据权利要求 16 所述的方法，所述方法还包括:

接收所述第一设备发送的所述第一数据以及所述第二数据是否匹配的匹配结果。

23、根据权利要求 16 所述的方法，其中，所述采用所述第一目标公钥对所述第二向量进行半同态加密生成第二加密向量包括:

针对所述第二向量中的每个第二分量，基于所述第一目标公钥对所述第二向量中的每个第二分量分别进行半同态加密，生成所述第二加密向量。

24、一种数据匹配装置，应用于第一设备，所述装置包括:

第一获取模块，用于将待匹配的第一数据输入到预先训练完成的向量转化模型中，获得所述第一数据对应的第一向量;

第一处理模块，用于采用自身生成的第一目标公钥对所述第一向量进行半同态加密生成第一加密向量，将所述第一目标公钥以及所述第一加密向量发送给第二设备;

所述第一获取模块，还用于获取所述第二设备发送的基于所述第一加密向量和第二加密向量计算的第一加密距离，所述第二加密向量为采用所述第一目标公钥对第二向量进行半同态加密后得到的；所述第二向量为将第二数据输入到所述第二设备中的预先训练完成的向量转化模型中获得的;

第一确定模块，用于基于所述第一加密距离以及与所述第一目标公钥对应的第一目标私钥，确定所述第一向量和第二向量的目标距离；基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否

匹配。

25、一种数据匹配装置，应用于第二设备，所述装置包括：

第二获取模块，用于将待匹配的第二数据输入到预先训练完成的向量转化模型中，获得所述第二数据对应的第二向量；

第二处理模块，用于接收第一设备发送的第一目标公钥以及第一加密向量，采用所述第一目标公钥对所述第二向量进行半同态加密生成第二加密向量；其中，所述第一加密向量为采用所述第一目标公钥对第一向量进行半同态加密后得到的，所述第一向量为将第一数据输入到所述第一设备中的预先训练完成的向量转化模型中获得的；

第二确定模块，用于基于所述第一加密向量和第二加密向量计算第一加密距离，将所述第一加密距离发送给所述第一设备，以使所述第一设备基于所述第一加密距离以及与所述第一目标公钥对应的第一目标私钥，确定所述第一向量和第二向量的目标距离；并使所述第一设备基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配。

26、一种数据匹配系统，所述系统包括：

第一设备，用于将待匹配的第一数据输入到预先训练完成的向量转化模型中，获得所述第一数据对应的第一向量；采用自身生成的第一目标公钥对所述第一向量进行半同态加密生成第一加密向量，将所述第一目标公钥以及所述第一加密向量发送给第二设备；

所述第二设备，用于将待匹配的第二数据输入到预先训练完成的向量转化模型中，获得所述第二数据对应的第二向量；并接收所述第一设备发送的第一目标公钥以及第一加密向量，采用所述第一目标公钥对所述第二向量进行半同态加密生成第二加密向量；基于所述第一加密向量和第二加密向量计算第一加密距离，将所述第一加密距离发送给所述第一设备；

所述第一设备，还用于基于所述第一加密距离以及与所述第一目标公钥对应的第一目标私钥，确定所述第一向量和第二向量的目标距离；基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是

否匹配。

27、一种电子设备，所述电子设备包括处理器和存储器，所述存储器用于存储程序指令，所述处理器用于执行存储器中存储的计算机程序时实现权利要求 1-23 任一所述数据匹配方法的步骤。

28、一种计算机可读存储介质，其存储有计算机程序，所述计算机程序被处理器执行时实现权利要求 1-23 任一所述数据匹配方法的步骤。

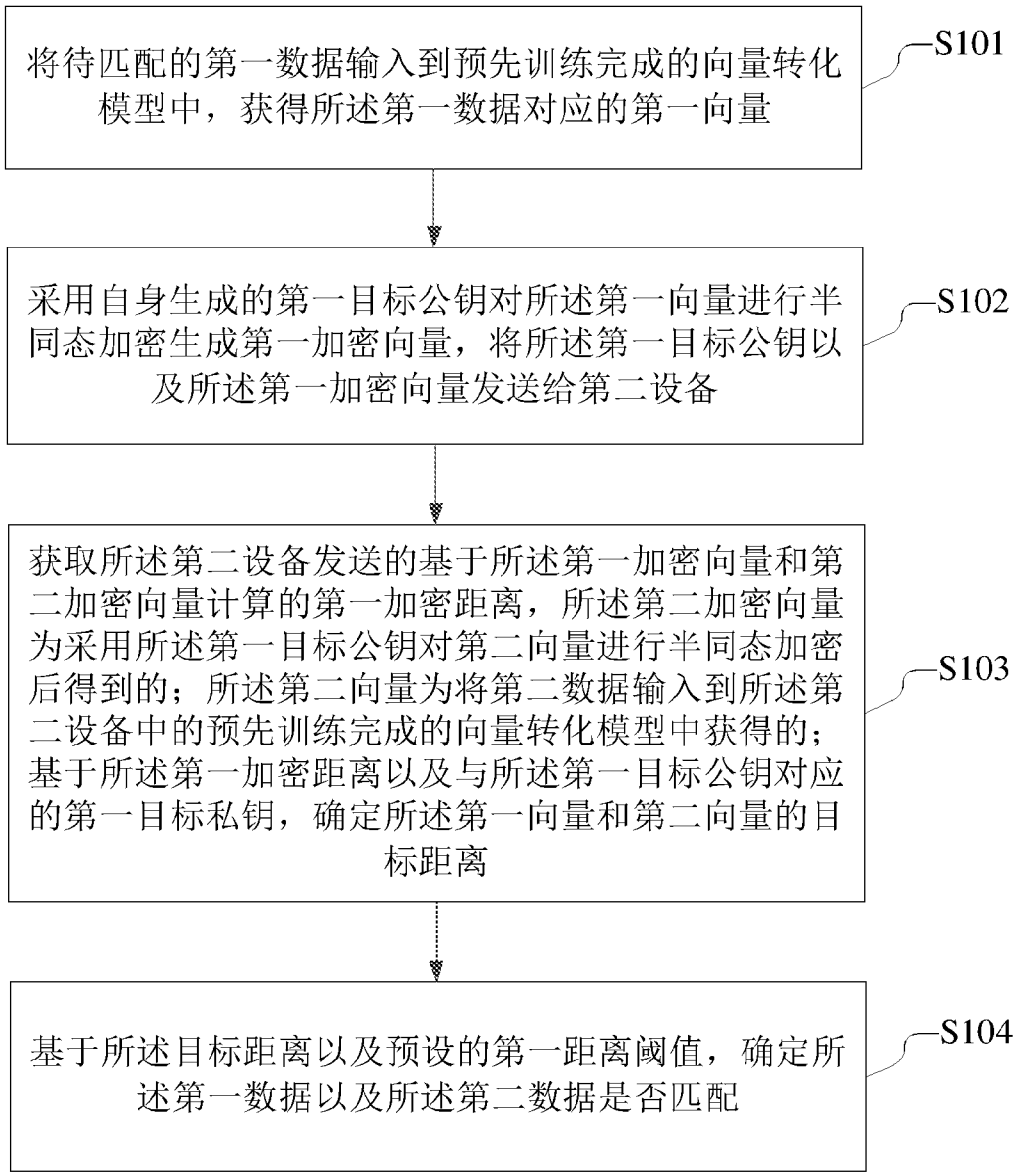

将待匹配的第一数据输入到预先训练完成的向量转化模型中，获得所述第一数据对应的第一向量
S101
采用自身生成的第一目标公钥对所述第一向量进行半同态加密生成第一加密向量，将所述第一目标公钥以及所述第一加密向量发送给第二设备
S102
获取所述第二设备发送的基于所述第一加密向量和第二加密向量计算的第一加密距离，所述第二加密向量为采用所述第一目标公钥对第二向量进行半同态加密后得到的；所述第二向量为将第二数据输入到所述第二设备中的预先训练完成的向量转化模型中获得的；基于所述第一加密距离以及与所述第一目标公钥对应的第一目标私钥，确定所述第一向量和第二向量的目标距离
S103
基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配
S104

图 1

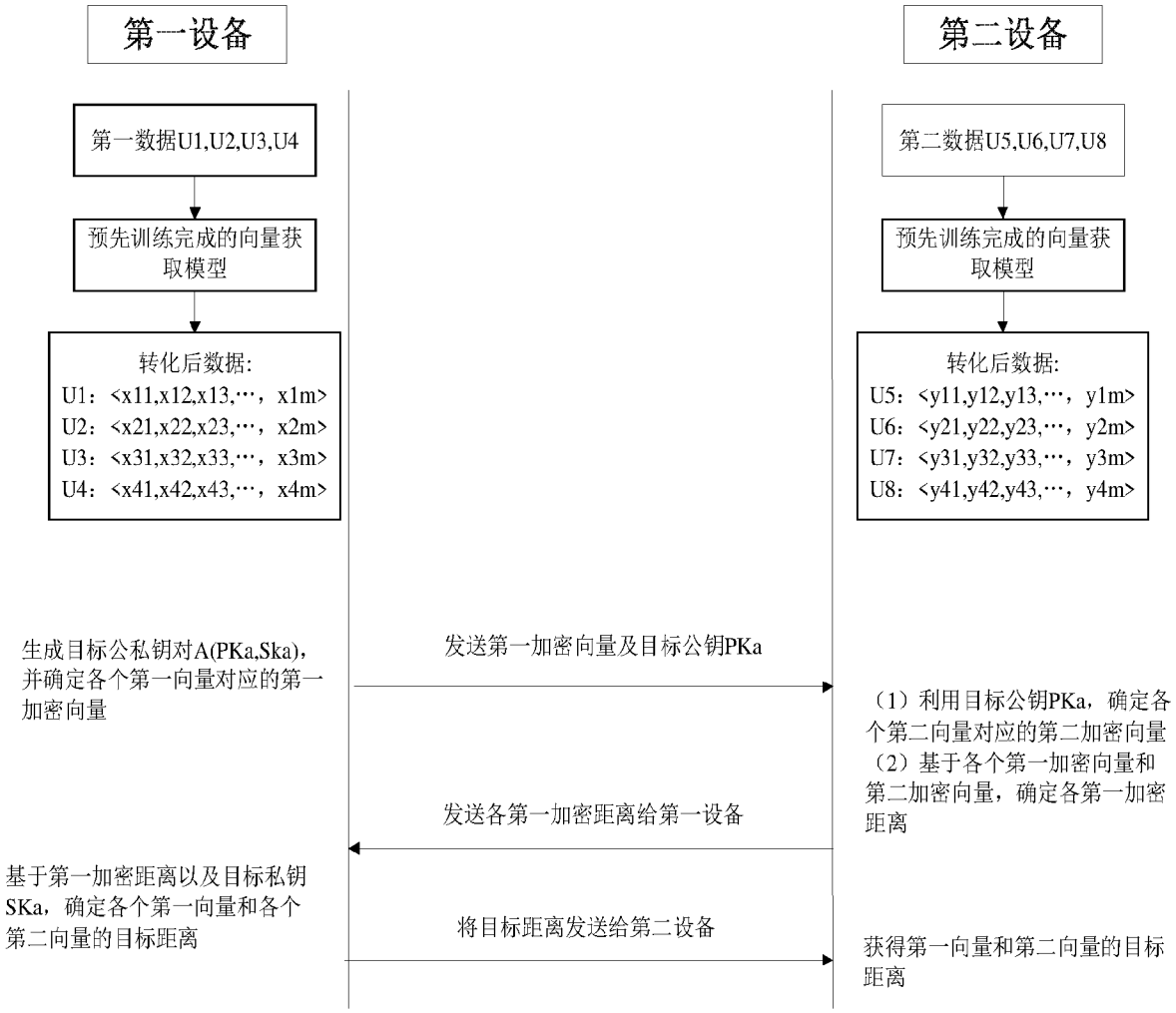

第一设备
第二设备
第一数据U1,U2,U3,U4
预先训练完成的向量获取模型
转化后数据:
U1：<x11,x12,x13,…，x1m>
U2：<x21,x22,x23,…，x2m>
U3：<x31,x32,x33,…，x3m>
U4：<x41,x42,x43,…，x4m>
第二数据U5,U6,U7,U8
预先训练完成的向量获取模型
转化后数据:
U5：<y11,y12,y13,…，y1m>
U6：<y21,y22,y23,…，y2m>
U7：<y31,y32,y33,…，y3m>
U8：<y41,y42,y43,…，y4m>
生成目标公私钥对A(PKa,Ska)，并确定各个第一向量对应的第一加密向量
发送第一加密向量及目标公钥PKa
（1）利用目标公钥PKa，确定各个第二向量对应的第二加密向量
（2）基于各个第一加密向量和第二加密向量，确定各第一加密距离
发送各第一加密距离给第一设备
基于第一加密距离以及目标私钥SKa，确定各个第一向量和各个第二向量的目标距离
将目标距离发送给第二设备
获得第一向量和第二向量的目标距离

图 2

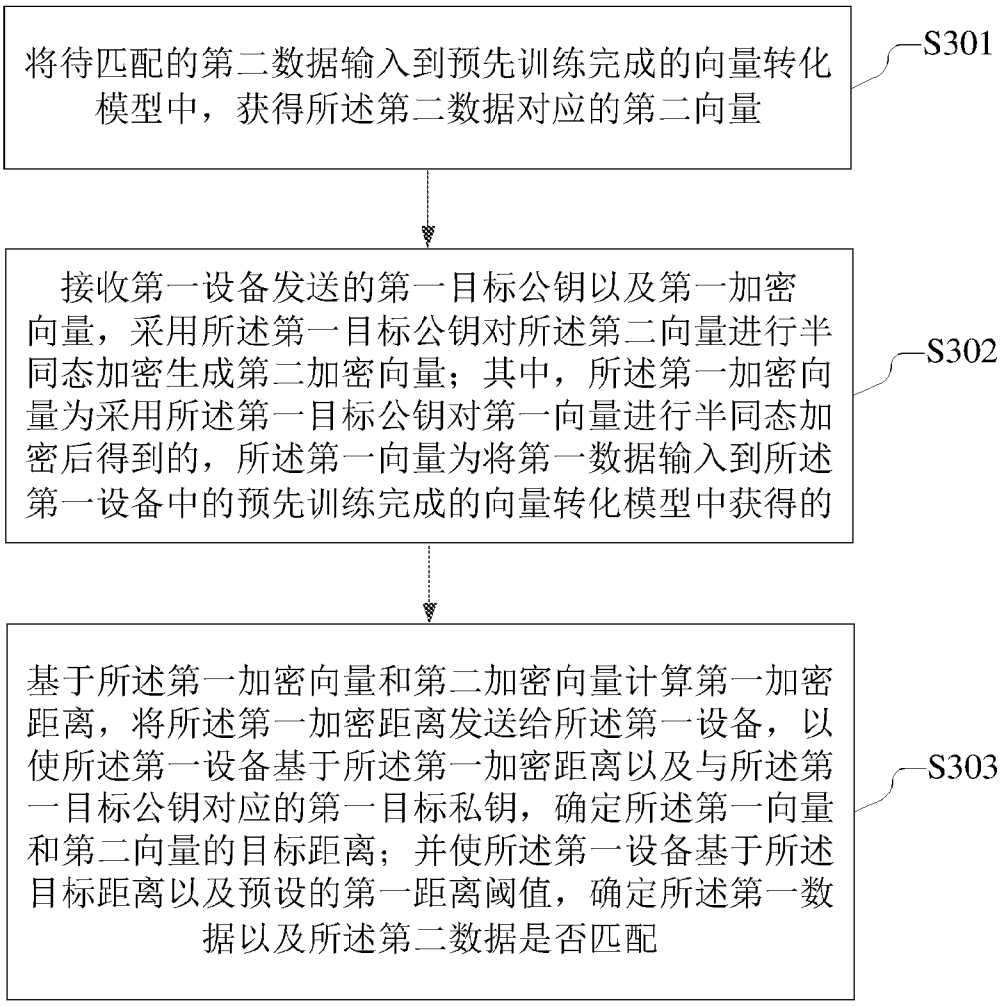

将待匹配的第二数据输入到预先训练完成的向量转化模型中，获得所述第二数据对应的第二向量
S301
接收第一设备发送的第一目标公钥以及第一加密向量，采用所述第一目标公钥对所述第二向量进行半同态加密生成第二加密向量；其中，所述第一加密向量为采用所述第一目标公钥对第一向量进行半同态加密后得到的，所述第一向量为将第一数据输入到所述第一设备中的预先训练完成的向量转化模型中获得的
S302
基于所述第一加密向量和第二加密向量计算第一加密距离，将所述第一加密距离发送给所述第一设备，以使所述第一设备基于所述第一加密距离以及与所述第一目标公钥对应的第一目标私钥，确定所述第一向量和第二向量的目标距离；并使所述第一设备基于所述目标距离以及预设的第一距离阈值，确定所述第一数据以及所述第二数据是否匹配
S303

图 3

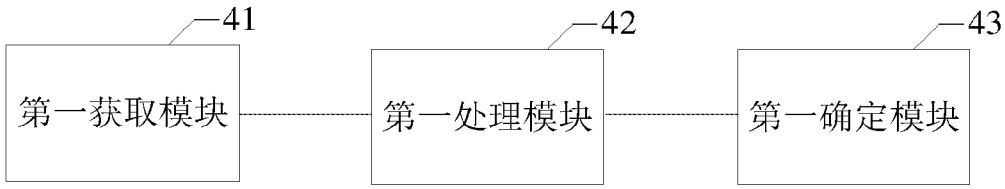

41
第一获取模块
42
第一处理模块
43
第一确定模块

图 4

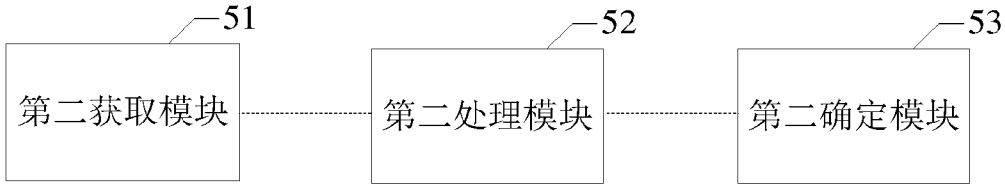

51
52
53
第二获取模块
第二处理模块
第二确定模块

图 5

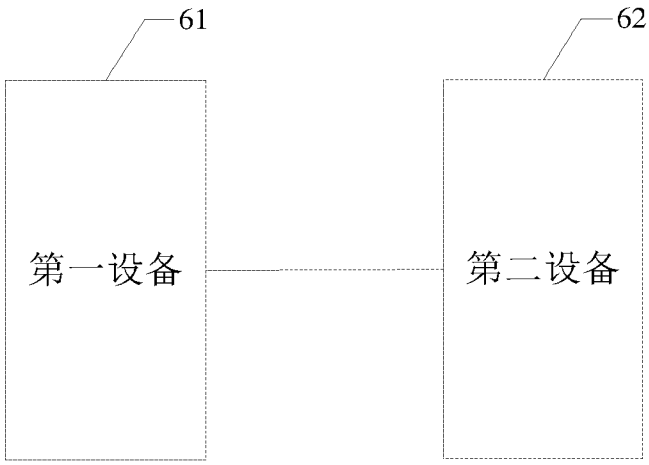

61
62
第一设备
第二设备

图 6

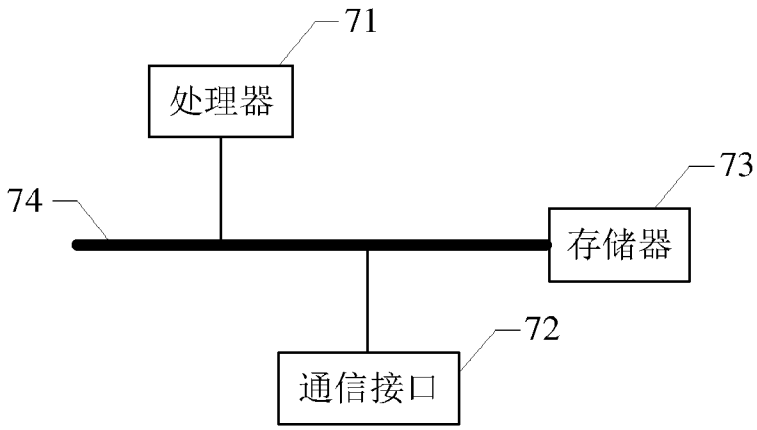

71
处理器
73
74
存储器
72
通信接口

图 7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2022/137361

A. CLASSIFICATION OF SUBJECT MATTER

H04L9/00(2022.01)i;G06F21/60(2013.01)n;H04L9/08(2006.01)n

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: H04L, G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNKI, CNTXT, ENTXT, VEN, IEEE: 半同态, 加密, 距离, 匹配, 向量, half homomorphic, part+ homomorphic, encrypt+, distance, match+, vector

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 115442021 A (CHINA UNIONPAY CO., LTD.) 06 December 2022 (2022-12-06) claims 1-28	1-28
Y	CN 114817943 A (CHINA UNIONPAY CO., LTD.) 29 July 2022 (2022-07-29) claims 1-32	1-28
Y	US 2018212759 A1 (ENVEIL, INC.) 26 July 2018 (2018-07-26) description, paragraph 0029	1-28
A	CN 112329519 A (ENGINEERING UNIVERSITY OF PAP) 05 February 2021 (2021-02-05) entire document	1-28
A	US 2015046708 A1 (FUJITSU LTD.) 12 February 2015 (2015-02-12) entire document	1-28
A	US 2016119119 A1 (XEROR CORP.) 28 April 2016 (2016-04-28) entire document	1-28

☐ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"D" document cited by the applicant in the international application
"E" earlier application or patent but published on or after the international filing date
"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)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"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
"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
"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
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
17 February 2023	**22 February 2023**

Name and mailing address of the ISA/CN	Authorized officer
China National Intellectual Property Administration (ISA/CN) **China No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088**	
Facsimile No. **(86-10)62019451**	Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2022/137361

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 115442021 A	06 December 2022	None	
CN 114817943 A	29 July 2022	None	
US 2018212759 A1	26 July 2018	None	
CN 112329519 A	05 February 2021	None	
US 2015046708 A1	12 February 2015	EP 2858296 A2	08 April 2015
		EP 2858296 A3	02 September 2015
		EP 2858296 B1	25 December 2019
		JP 2015031935 A	16 February 2015
		JP 6244728 B2	13 December 2017
		US 9509493 B2	29 November 2016
US 2016119119 A1	28 April 2016	US 9749128 B2	29 August 2017
		DE 102015208142 A1	19 November 2015
		JP 2016035554 A	17 March 2016

国际检索报告

国际申请号
PCT/CN2022/137361

A. 主题的分类

H04L9/00(2022.01)i;G06F21/60(2013.01)n;H04L9/08(2006.01)n

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: H04L, 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））

CNKI,CNTXT,ENTXT,VEN,IEEE:半同态, 加密, 距离, 匹配, 向量, half homomorphic, part+ homomorphic, encr-ypt+, distance, match+, vector

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 115442021 A (中国银联股份有限公司) 2022年12月6日 (2022-12-06) 权利要求1-28	1-28
Y	CN 114817943 A (中国银联股份有限公司) 2022年7月29日 (2022-07-29) 权利要求1-32	1-28
Y	US 2018212759 A1 (ENVEIL INC.) 2018年7月26日 (2018-07-26) 说明书第0029段	1-28
A	CN 112329519 A (中国人民武装警察部队工程大学) 2021年2月5日 (2021-02-05) 全文	1-28
A	US 2015046708 A1 (FUJITSU LTD.) 2015年2月12日 (2015-02-12) 全文	1-28
A	US 2016119119 A1 (XEROX CORP.) 2016年4月28日 (2016-04-28) 全文	1-28

☐ 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	国际检索报告邮寄日期
2023年2月17日	2023年2月22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 传真号 (86-10)62019451	受权官员 傅琦 电话号码 (+86) 010-62411842

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2022/137361

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	115442021	A	2022年12月6日	无			
CN	114817943	A	2022年7月29日	无			
US	2018212759	A1	2018年7月26日	无			
CN	112329519	A	2021年2月5日	无			
US	2015046708	A1	2015年2月12日	EP	2858296	A2	2015年4月8日
				EP	2858296	A3	2015年9月2日
				EP	2858296	B1	2019年12月25日
				JP	2015031935	A	2015年2月16日
				JP	6244728	B2	2017年12月13日
				US	9509493	B2	2016年11月29日
US	2016119119	A1	2016年4月28日	US	9749128	B2	2017年8月29日
				DE	102015208142	A1	2015年11月19日
				JP	2016035554	A	2016年3月17日