

DATA STATISTICS METHOD AND APPARATUS

Technical Field

5 The present invention relates to the network technology field, and more particularly to a data statistics method and apparatus.

Background Art

A great number of data islands exist in the big data era. For example, a natural person's data can be dispersedly stored in different enterprises. However, there is no complete mutual trust between enterprises due to the competitive relationship and the consideration of the user privacy protection, which creates a barrier for statistical tasks of data cooperation between enterprises. There is an urgent need for a solution on how to complete some data statistics calculations using the data owned by two parties without leaking data privacy of the respective enterprises, and to ensure that the core data privacy of the respective enterprises is fully protected. However, no good solution is available now.

15 Summary of the Invention

For this reason, the present invention provides a data statistics method and apparatus to implement secure calculation between two parties while protecting data privacy of the two data owners.

20 Specifically, one or more embodiments of the description are implemented by the following technical solution.

In a first aspect, a data statistics method is provided. The method may be applied to data statistics by combining data of a local data party and a cooperative data party, the local data party having a plurality of pieces of first data for calculating statistical values, the plurality of pieces of first data respectively corresponding to different data identifiers, and the cooperative data party 25 having a plurality of pieces of second data corresponding to the data identifiers. The method comprises:

respectively performing homomorphic encryption on the plurality of pieces of first data to obtain encrypted data;

30 sending the data identifiers corresponding to the plurality of pieces of first data and the encrypted data to the cooperative data party;

receiving encrypted statistical values returned by the cooperative data party, the encrypted statistical values being obtained by the cooperative data party by performing a statistical calculation on encrypted data corresponding to various data identifiers in an identifier intersection, the identifier intersection being identifiers corresponding to pieces of second data of the cooperative data party for the data statistics and selected from the plurality of data identifiers corresponding to the plurality of pieces of first data; and

performing homomorphic decryption on the encrypted statistical values and obtaining the statistical values.

In a second aspect, a data statistics method is provided, the method being applied to data statistics by combining data of a local data party and a statistical data party, the statistical data party having a plurality of pieces of first data for calculating statistical values, the plurality of pieces of first data respectively corresponding to different data identifiers, and the local data party having a plurality of pieces of second data corresponding to the data identifiers. The method comprises:

receiving data identifiers and corresponding encrypted data from the statistical data party, the data identifiers being identifiers corresponding to pieces of first data of the statistical data party for the data statistics, and the encrypted data being obtained by the statistical data party by performing homomorphic encryption on the pieces of first data;

determining an identifier intersection according to data identifiers corresponding to pieces of second data of the local data party for the data statistics and the data identifiers of the pieces of first data;

performing statistical processing on encrypted data corresponding to various data identifiers in the identifier intersection to obtain encrypted statistical values; and

sending the encrypted statistical values to the statistical data party to enable the statistical data party to perform homomorphic decryption on the encrypted statistical values to obtain the statistical values.

In a third aspect, a data statistics method is provided, the method being used for data statistics between a local data party and a cooperative data party, the local data party having first data corresponding to data identifiers, the cooperative data party storing second data corresponding to the same data identifiers, and the method being applied to obtain statistical values from a plurality of pieces of first data. The method comprises:

performing homomorphic encryption on the plurality of pieces of first data of the local data party for the data statistics to obtain encrypted data;

performing local private key processing on data identifiers respectively corresponding to the plurality of pieces of first data according to a key exchange protocol to obtain local processing identifiers;

sending the local processing identifiers and the encrypted data corresponding to the pieces of first data to the cooperative data party, to enable the cooperative data party to generate first key processing identifiers by performing peer private key processing on the local processing identifiers, and store corresponding relations between the first key processing identifiers and the encrypted data;

receiving second processing identifiers from the cooperative data party, the second processing identifiers being obtained by the cooperative data party by performing the peer private key processing on data identifiers corresponding to second data for the data statistics;

performing the local private key processing on the second processing identifiers to generate second key processing identifiers, and sending the second key processing identifiers to the cooperative data party;

receiving encrypted statistical values from the cooperative data party, the encrypted statistical values being obtained by the cooperative data party by performing statistical processing on various encrypted data corresponding to an identifier intersection between the first key processing identifiers and the second key processing identifiers; and

performing homomorphic decryption on the encrypted statistical values and obtaining the statistical values.

In a fourth aspect, a data statistics method is provided, the method being used for data statistics between a local data party and a statistical data party, the statistical data party having first data corresponding to data identifiers, the local data party storing second data corresponding to the data identifiers, and the method being applied to obtain statistical values from a plurality of pieces of first data. The method comprises:

receiving a plurality of peer processing identifiers and a plurality of pieces of encrypted data respectively corresponding to the plurality of peer processing identifiers from the statistical data party, the plurality of peer processing identifiers being obtained by the statistical data party by performing peer private key processing on data identifiers of the plurality of pieces of first data

for the data statistics according to a key exchange protocol, the plurality of pieces of encrypted data being obtained by the statistical data party by performing homomorphic encryption on the first data;

performing local private key processing on the plurality of peer processing identifiers according to a key exchange protocol for generating first key processing identifiers, and storing corresponding relations between the first key processing identifiers and the corresponding encrypted data;

performing the local private key processing on data identifiers respectively corresponding to a plurality of pieces of second data of the local data party for the data statistics according to the key exchange protocol to obtain second processing identifiers;

sending the second processing identifiers to the statistical data party, and receiving second key processing identifiers returned by the statistical data party, the second key processing identifiers being obtained by the statistical data party by performing the peer private key processing on the second processing identifiers;

obtaining various encrypted data corresponding to an identifier intersection between the first key processing identifiers and the second key processing identifiers, and performing statistical processing on the various encrypted data to obtain encrypted statistical values; and

sending the encrypted statistical values to the statistical data party for the statistical data party to perform homomorphic decryption on the encrypted statistical values to obtain the statistical values.

In a fifth aspect, a data statistics apparatus is provided, the apparatus being used for data statistics between a local data party and a cooperative data party, the local data party having first data corresponding to data identifiers, the cooperative data party having second data corresponding to the data identifiers, and the apparatus being applied to obtain statistical values from a plurality of pieces of first data. The apparatus comprises:

a data encryption module that is used to respectively perform homomorphic encryption on the plurality of pieces of first data to obtain encrypted data;

a data sending module that is used to send data identifiers corresponding to the plurality of pieces of first data and the encrypted data to the cooperative data party;

a data receiving module that is used to receive encrypted statistical values returned by the cooperative data party, the encrypted statistical values being obtained by the cooperative data

party by performing a statistical calculation on encrypted data corresponding to various data identifiers in an identifier intersection, the identifier intersection being identifiers corresponding to pieces of second data of the cooperative data party for the data statistics and selected from the plurality of data identifiers corresponding to the plurality of pieces of first data; and

5 a decryption processing module that is used to perform homomorphic decryption on the encrypted statistical values and obtain the statistical values.

In a sixth aspect, a data statistics apparatus is provided, the apparatus being used for data statistics between a local data party and a statistical data party, the statistical data party having a plurality of pieces of first data for calculating statistical values, the plurality of pieces of first data respectively corresponding to different data identifiers, and the local data party having a
10 plurality of pieces of second data corresponding to the data identifiers. The apparatus comprises:

a data obtaining module that is used to receive data identifiers and corresponding encrypted data from the statistical data party, the data identifiers being identifiers corresponding to pieces of first data of the statistical data party for the data statistics, and the encrypted data being
15 obtained by the statistical data party by performing homomorphic encryption on the plurality of pieces of first data;

an intersection determining module that is used to determine an identifier intersection according to data identifiers corresponding to pieces of second data of the local data party for the data statistics and the data identifiers of the pieces of first data;

20 a statistical processing module that is used to perform statistical processing on encrypted data corresponding to various data identifiers in the identifier intersection to obtain encrypted statistical values; and

a statistical sending module that is used to send the encrypted statistical values to the statistical data party to enable the statistical data party to perform homomorphic decryption on
25 the encrypted statistical values to obtain the statistical values.

In a seventh aspect, a data statistics apparatus is provided, the apparatus comprising a memory, a processor, and computer instructions stored on the memory and running on the processor, the processor implementing the following steps during the execution of instructions:

respectively performing homomorphic encryption on the plurality of pieces of first data to
30 obtain encrypted data;

sending the data identifiers corresponding to the plurality of pieces of first data and the encrypted data to the cooperative data party;

receiving encrypted statistical values returned by the cooperative data party, the encrypted statistical values being obtained by the cooperative data party by performing statistical

5 calculation on encrypted data corresponding to various data identifiers in an identifier intersection, the identifier intersection being identifiers corresponding to pieces of second data of the cooperative data party for the data statistics and selected from the plurality of data identifiers corresponding to the plurality of pieces of first data; and

performing homomorphic decryption on the encrypted statistical values to obtain the
10 statistical values.

In an eighth aspect, a data statistics apparatus is provided, the apparatus being used for data statistics between a local data party and a statistical data party, the statistical data party storing first data corresponding to data identifiers, the local data party storing second data corresponding to the data identifiers, and the apparatus being applied to obtain statistical values from a plurality
15 of pieces of first data. The apparatus comprises:

a data receiving module that is used to receive a plurality of peer processing identifiers and a plurality of pieces of encrypted data respectively corresponding to the plurality of peer processing identifiers from the statistical data party, the plurality of peer processing identifiers being obtained by the statistical data party by performing peer private key processing on data
20 identifiers of the plurality of pieces of first data for the data statistics according to a key exchange protocol, the plurality of pieces of encrypted data being obtained by the statistical data party by performing homomorphic encryption on the first data;

a relation storage module that is used to perform local private key processing on the plurality of peer processing identifiers according to a key exchange protocol to generate first key
25 processing identifiers, and store corresponding relations between the first key processing identifiers and the corresponding encrypted data;

a private key encryption module that is used to perform the local private key processing on data identifiers respectively corresponding to a plurality of pieces of second data of the local data party for the data statistics according to the key exchange protocol to obtain second processing
30 identifiers;

an identifier sending module that is used to send the second processing identifiers to the statistical data party, and receive second key processing identifiers returned by the statistical data party, the second key processing identifiers being obtained by the statistical data party by performing the peer private key processing on the second processing identifiers;

5 a statistical processing module that is used to obtain various encrypted data corresponding to an identifier intersection between the first key processing identifiers and the second key processing identifiers, and perform statistical processing on the various encrypted data to obtain encrypted statistical values; and

10 a data sending module that is used to send the encrypted statistical values to the statistical data party for the statistical data party to perform homomorphic decryption on the encrypted statistical values and obtain the statistical values.

In a ninth aspect, a data statistics apparatus is provided, the apparatus being used for data statistics between a local data party and a cooperative data party, the local data party having first data corresponding to data identifiers, the cooperative data party storing second data
15 corresponding to the same data identifiers; in addition, the apparatus being applied to obtaining statistical values from a plurality of pieces of first data; the apparatus comprising:

 a data encryption module that is used to perform homomorphic encryption on the plurality of pieces of first data of the local data party for the data statistics to obtain encrypted data;

20 an identifier processing module that is used to perform local private key processing on data identifiers respectively corresponding to the plurality of pieces of first data according to a key exchange protocol to obtain a plurality of local processing identifiers;

25 a data sending module that is used to send the plurality of local processing identifiers and the encrypted data to the cooperative data party to enable the cooperative data party to generate first key processing identifiers by performing peer private key processing on local processing identifiers and store corresponding relations between the first key processing identifiers and the encrypted data;

30 an identifier receiving module that is used to receive second processing identifiers from the cooperative data party, the second processing identifiers being obtained by the cooperative data party by performing the peer private key processing on data identifiers corresponding second data for the data statistics;

a private key processing module that is used to perform the local private key processing on the second processing identifiers to generate and send second key processing identifiers to the cooperative data party;

5 a statistical receiving module that is used to receive encrypted statistical values from the cooperative data party, the encrypted statistical values being obtained by the cooperative data party by performing statistical processing on various encrypted data corresponding to an intersection between the first key processing identifiers and the second key processing identifiers;

10 a data decryption module that is used to perform homomorphic decryption on the encrypted statistical values and obtain statistical values.

In the data statistics method and apparatus according to one or more embodiments of the description, homomorphically encrypted data is sent to a cooperative data party so that cooperative data party can perform statistical processing on the encrypted data. The same statistical processing results can be obtained as long as local homomorphic decryption is
15 performed on encrypted statistical values. This method protects the local data privacy security, enables processing the combined data of two data sources without exposing any data privacy, and implements secure calculation between two parties while protecting data privacy of two data owners.

Brief Description of the Figures

20 To clearly describe technical solutions in one or more embodiments of the description or in the prior art, figures to be used in the description of the embodiments or of the prior art will be briefly introduced below. Obviously, the figures in the following description are only some embodiments recorded in one or more embodiments of the description. Ordinary technicians in this field may also obtain other figures according to these figures without creative efforts.

25 FIG. 1 is a flow chart of a data statistics method according to one or more embodiments of the description.

FIG. 2 is a flow chart of a data statistics method according to one or more embodiments of the description.

30 FIG. 3 is a structural schematic diagram of a data statistics apparatus according to one or more embodiments of the description.

FIG. 4 is a structural schematic diagram of a data statistics apparatus according to one or more embodiments of the description.

FIG. 5 is a structural schematic diagram of a data statistics apparatus according to one or more embodiments of the description.

5 FIG. 6 is a structural schematic diagram of a data statistics apparatus according to one or more embodiments of the description.

Particular Embodiments

10 A clear and complete description of the technical solutions in one or more embodiments of the description will be provided below in conjunction with the figures in one or more embodiments of the description so that people in this technical field can understand the technical solutions in one or more embodiments of the description better. Obviously, the embodiments described are only a part of the embodiments and are not all embodiments. All other embodiments obtained by ordinary technicians in the field based on one or more embodiments of the description without creative efforts also fall in the scope of protection of the present invention.

15 The data storage method may be a vertical mode in the big data era, that is, a plurality of data owners may own different attribute information of the same entity. For example, a person's vehicle insurance score exists in one organization while the same person's claim amount exists in another organization. Such data storage in vertical mode may cause some data statistics calculations involving a plurality of data owners, and require that a plurality of data owners cooperates to complete one data statistics. However, data confidentiality of respective enterprises cannot be breached due to the competitive relationship between different enterprises and the concern of privacy protection.

20 In the examples of the present invention, the purpose is to perform data statistics based on different data owners' data without disclosing data owners' respective data privacy. This method is elaborated below using an exemplary application scenario but is not limited to this application scenario.

Application scenario:

30 Taking the data statistics of the insurance score of the vehicle insurance as an example, there are two data sources in this example: data sources A and B. Assume that the data source A

is a data organization, and the data source B is an insurance organization. Both data sources respectively store different information of the same vehicle owner.

Data source A: Assume that this data source A stores every vehicle owner's vehicle insurance score. A vehicle insurance score is a score obtained after accurately portraying and risk analysis is performed on a vehicle owner. A higher vehicle insurance score indicates a lower risk. As shown in Table 1, the data structure of vehicle insurance scores stored in the data source A is as follows:

Table 1 Data structure of data source A

Column Name	Type	Description	Example
idcard_no	string	ID card No.	*****197309119564
score	int	Vehicle insurance score	510

Data source B: Assume that this data source B stores every vehicle owner's claim information. For example, a vehicle owner's claim information includes the number of claims, the claim amount, etc. As shown in Table 2, data structure of every vehicle owner stored on the data source B side is as follows:

Table 2 Data structure of data source B

Column Name	Type	Description	Example
idcard_no	string	ID card No.	*****197309119564
gender	string	Gender	female
times	int	Number of claims in the recent one year	3
amount	int	Claim amount	3,500

Based on the preceding application scenario, one data statistics can be jointly finished based on the data of the data sources A and B. For example, the required statistical task is "to perform statistical calculation on the sum of vehicle insurance scores of female users whose number of claims is greater than 5". In this required task, "female users whose number of claims is greater than 5" indicates that data in the data source B can be used as a filtering condition for obtaining a statistical value. "Sum of vehicle insurance scores" indicates that the summation statistic calculation of the data "score" in the data source A can be performed.

Based on the data structure as shown in Table 1, assume that the data source A owns data as shown in the following Table 3, wherein idcard_no is a vehicle owner's ID card number and score is this vehicle owner's vehicle insurance score.

Table 3 Data of data source A

idcard_no	score
1234567	490
2345678	501
3456789	530

5 Based on the data structure as shown in Table 2, assume that the data source B owns data as shown in the following Table 4:

Table 4 Data of data source B

idcard_no	gender	times	amount
1234567	Male	3	5000
2345678	Female	7	23000
3456789	Female	6	16000

10 Assume that the statistical calculation on the sum of vehicle insurance scores of female users whose number of claims is greater than 5 will be performed based on the preceding Tables 3 and 4. It can be seen that the statistical data of "vehicle insurance score" in this statistics is stored in the data source A. The score column in Table 3 is called "statistical column", that is, statistical calculation of data in this column will be performed to sum the vehicle insurance scores. "Number of claims" and "female" in the filtering conditions are stored in the data source 15 B. Therefore, the cooperation between the data sources A and B is required to finish the summation statistical calculation of vehicle insurance scores.

In the following description of the data statistics method, the data source A with the statistical data "vehicle insurance score" is called statistical data party, while the other data source B is called cooperative data party. As mentioned above, both data sources respectively 20 store different information of the same vehicle owner. The vehicle owner's information stored in the data source A (e.g. score) is called first data. The vehicle owner's information about parameter data statistics stored in the data source B (e.g. gender, number of claims in the recent one year, and claim amount) is called second data. The second data is used as filtering conditions

during the statistical calculation. In addition, idcard_no included in both the data sources A and B is called data identifier, that is, the data source A stores the first data corresponding to this data identifier and the data source B stores the second data corresponding to the same data identifier.

FIG. 1 shows an example of a process of a data statistics method. As shown in FIG.1, the method comprises the following steps.

In Step 100, the statistical data party respectively performs homomorphic encryption on the plurality of pieces of first data and obtains encrypted data.

The plurality of pieces of first data in this step is the data of statistical data party for the data statistics. The data is obtained according to the data filtering conditions of the statistical data party.

In Step 102, the statistical data party sends data identifiers corresponding to the plurality of pieces of first data and the encrypted data to a cooperative data party.

In Step 104, the cooperative data party determines an identifier intersection according to data identifiers corresponding to a plurality of pieces of second data of the local data party for the data statistics as well as the data identifiers of the plurality of pieces of first data.

In this step, the cooperative data party obtains the second data of this data statistics according to the local filtering conditions and acquires the data identifiers corresponding to the second data. The intersection between the data identifiers acquired by the cooperative data party and the received data identifiers from the statistical data party in Step 102 is called the identifier intersection. The identifier intersection includes at least one data identifier. The first data corresponding to various data identifiers in the identifier intersection is the data of the statistical data party for the data statistics. The second data corresponding to the data identifiers is the data of the cooperative data party for the data statistics.

In Step 106, the cooperative data party performs statistical processing on the encrypted data corresponding to various data identifiers in the identifier intersection and obtains encrypted statistical values thereof.

In this step, the cooperative data party performs statistical calculation on the encrypted data corresponding to various data identifiers in the identifier intersection. Specific types of the data statistics, such as summation and averaging, depend on specific business requirements. In addition, the statistical processing here is performed on the encrypted data, not on the original first data. This statistical method will not expose true data of the statistical data party, and allows

the cooperative data party to freely perform the required data statistics satisfying its own data filtering conditions.

In Step 108, the cooperative data party sends the encrypted statistical values to the statistical data party.

5 In Step 110, the statistical data party performs homomorphic decryption on the encrypted statistical values and obtains statistical values.

In this step, the statistical values that are the same as those obtained from the original first data statistics can be obtained by performing the homomorphic decryption on the returned encrypted statistical values according to the characteristics of homomorphic encryption.

10 The data statistics method in this example is applied to calculation of two parties' combined data by the homomorphic encryption method so that the data received by the cooperative data party is the homomorphically encrypted data without exposing the statistical data party's true data. In addition, when receiving the encrypted statistical values, the statistical data party can obtain the statistical values by performing the homomorphic decryption. This is very convenient
15 and also protects both data owners' data privacy.

In another example, when data is transmitted between a local data party and a cooperative data party, the data identifiers are encrypted according to a key exchange protocol to provide more secure protection for both data owners' data privacy. For example, the statistical data party performs local private key processing on the data identifiers using a local private key, and sends
20 the processed data identifiers to the peer so that the peer can continue to perform peer private key processing on the data identifiers. The statistical data party also receives the data identifiers sent by the cooperative data party that are processed by the peer key processing, and continues to perform the local private key processing on the processed data identifiers before returning them to the cooperative data party. The processing of the data identifiers by both parties according to a
25 key exchange protocol avoids the exposure of the data identifiers and provides more secure protection.

The following FIG. 2 shows an example of a process of a data statistics method. This process can perform statistical calculations on the sum of vehicle insurance scores of female users whose number of claims is greater than 5 based on Tables 3 and 4. In addition, this
30 example takes the combination of the homomorphic encryption and key exchange processing as an example. As shown in FIG. 2, this method may comprise the following steps.

In Step 200, the statistical data party performs the homomorphic encryption on the plurality of pieces of first data of local data statistics, and obtains the encrypted data corresponding to the first data.

5 In this example, the data source A is a statistical data party. The score column is called a statistical column, and every vehicle score is called the first data.

In one example, the data source A performs statistical calculations of sum of all vehicle insurance scores in the score statistical column of the statistical Table 3, that is, the sum of three vehicle insurance scores including 490, 501, and 530. The three vehicle insurance scores, including 490, 501, and 530, are called "three pieces of first data of local data statistics".

10 In another example, the data source A can also select a part of vehicle insurance scores to participate in the summation statistical calculation according to the predetermined data filtering conditions. For example, the data source A selects two vehicle insurance scores including 501 and 530 to perform the summation.

15 In this step, assume that there is no data filtering condition on the data source A. The encrypted data respectively corresponding to various first data is obtained by performing the homomorphic encryption on all of the first data in the statistical column. For example, the encrypted data of the vehicle insurance scores of 490 and 501 is expressed as Enc (490) and Enc (501), respectively.

20 In Step 202, the statistical data party performs the local private key processing on the data identifiers respectively corresponding to the plurality of pieces of first data of the local data statistics according to the key exchange protocol, and obtains the plurality of first processing identifiers.

25 In this step, idcard_no corresponding to the score in Table 3 is called the data identifier corresponding to the first data. A key exchange protocol such as Diffie-Hellman key exchange ("D-H" for short) is used to process the preceding data identifiers to protect the detailed data of the data sources A and B from being leaked.

30 For example, Hash is performed on idcard_no to obtain H(K). At the same time, the data source A generates its own private key α in the key exchange protocol and performs the local private key processing. This processing is an α exponent operation on H(k) to obtain $H(k)^\alpha$. This $H(k)^\alpha$ is called a first processing identifier.

Taking all first data in the statistical column of the data source A participating in this statistical calculation of summation as an example, after processing of this step is performed, the data source A obtains the encrypted data corresponding to the various first data participating in the statistics and the first processing identifiers. As shown in the following Table 5, $H(k)^a$ (that is, $\text{Hash}(\text{idcard_no})^a$) is the first processing identifier and $\text{Enc}(\text{score})$ is the encrypted data. Taking the first vehicle owner in Table 3 as an example, the vehicle insurance score of this vehicle owner is 490, the corresponding encrypted data is $\text{Enc}(490)$, and the data identifier corresponding to this vehicle insurance score of 490 is 1234567. The first processing identifier $H(1234567)^a$ is obtained after Hash and the local private key processing are performed on this data identifier.

Table 5 First processing identifiers and encrypted data

$\text{Hash}(\text{idcard_no})^a$	$\text{Enc}(\text{score})$
$H(1234567)^a$	$\text{Enc}(490)$
$H(2345678)^a$	$\text{Enc}(501)$
$H(3456789)^a$	$\text{Enc}(530)$

In Step 204, the statistical data party sends first processing identifiers respectively corresponding to the plurality of pieces of first data and the encrypted data to the cooperative data party.

In this step, the data source A sends the data in Table 5 to the data source B.

In Step 206, the cooperative data party generates the first key processing identifiers, and stores the corresponding relations between the first key processing identifiers and the encrypted data after performing the local private key processing of the data source B on the first processing identifiers according to a key exchange protocol.

In this step, after receiving the data in Table 5, the data source B generates a local private key β of the data source B according to a key exchange protocol, and uses this private key β to perform the local private key operations (that is, exponent operations) on the first processing identifier $H(k)^a$ to obtain $H(k)^{a\beta}$. This $H(k)^{a\beta}$ is called the first key processing identifier. After processing of the β exponent operation of this step is performed, Table 5 is converted into Table 6 as follows.

Table 6 First key processing identifiers and encrypted data

$\text{Hash}(\text{idcard_no})^{\alpha\beta}$	Enc (score)
$H(1234567)^{\alpha\beta}$	Enc (490)
$H(2345678)^{\alpha\beta}$	Enc (501)
$H(3456789)^{\alpha\beta}$	Enc (530)

In Step 208, the cooperative data party performs the local private key processing on the data identifiers respectively corresponding to the plurality of pieces of second data of the local data statistics according to the key exchange protocol, and obtains a plurality of second processing identifiers.

In this step, the data source B also determines the plurality of pieces of second data of the local data statistics. For example, the plurality of pieces of second data may include all of the data or the filtered data obtained by perform local filtering according to the predetermined filtering conditions.

For example, the predetermined filtering condition is "female users whose number of claims is greater than 5". Data in Table 4 is filtered according to this condition to obtain the result that the last two rows of the data will participate in the statistics. Gender and number of claims of the data source B in Table 4 are called the second data, for example, "female, 7" and "female, 6". Data identifiers corresponding to the two pieces of second data are 2345678 and 3456789, respectively.

The data source B respectively performs Hash on the preceding data identifiers to obtain $H(K)$, and then perform the β exponent operation on $H(k)$ according to the key exchange protocol to obtain $H(k)^\beta$. This β is the private key of the data source B. This $H(k)^\beta$ is called the second processing identifier, as shown in the following Table 7.

Table 7 Second key processing identifiers

$\text{Hash}(\text{idcard_no})^\beta$
$H(2345678)^\beta$
$H(3456789)^\beta$

In Step 210, the cooperative data party sends the second processing identifiers to the statistical data party.

In this step, the data source B sends the data in the preceding Table 7 to the data source A.

In Step 212, the statistical data party generates the second key processing identifiers after performing the local private key processing on the second processing identifiers.

For example, after receiving $\text{Hash}(\text{idcard_no})^\beta$ in Table 7, the data source A generates a second key processing identifier $\text{Hash}(\text{idcard_no})^{\beta\alpha}$ by performing the local private key processing of the data source A, as shown in the following Table 8.

Table 8 Second key processing identifiers

$\text{Hash}(\text{idcard_no})^{\beta\alpha}$
$H(2345678)^{\beta\alpha}$
$H(3456789)^{\beta\alpha}$

In Step 214, the statistical data party sends the second key processing identifiers to the cooperative data party.

In Step 216, the cooperative data party obtains various encrypted data corresponding to the identifier intersection between the first and the second key processing identifiers, performs the statistical processing on the encrypted data, and obtains the encrypted statistical values.

In this step, the data source B gets an intersection between the second key processing identifiers in Table 8 and the first key processing identifiers in Table 6. $\text{Hash}(\text{idcard_no})^{\beta\alpha}$ and $\text{Hash}(\text{idcard_no})^{\alpha\beta}$ with the same value indicate that they correspond to the same idcard_no, that is, a vehicle owner represented by this idcard_no satisfies the filtering conditions of data of both the statistical data party and the cooperative data party participating in the statistics. The encrypted data corresponding to the first key processing identifiers in the intersection is obtained according to the intersection and the corresponding relations between the first key processing identifiers and the encrypted data in Table 6. As shown in the following Table 9, the intersection and the various encrypted data corresponding to the intersection are included in Table 9.

Table 9 Identifier intersection and corresponding encrypted data

$\text{Hash}(\text{idcard_no})^{\alpha\beta}$	$\text{Hash}(\text{idcard_no})^{\beta\alpha}$	Enc (score)
$H(2345678)^{\alpha\beta}$	$H(2345678)^{\beta\alpha}$	Enc (501)
$H(3456789)^{\alpha\beta}$	$H(3456789)^{\beta\alpha}$	Enc (530)

Statistical summation of the various corresponding encrypted data in the preceding identifier intersection can be performed according to Table 9, for example, $\text{Enc}(501) + \text{Enc}(530)$. Values obtained are called the encrypted statistical values.

In Step 218, the cooperative data party sends the encrypted statistical values to the statistical data party.

For example, the data source B sends the preceding encrypted statistical value "Enc (501)+Enc (530)" to the data source A.

5 In Step 220, the statistical data party performs the homomorphic decryption on the encrypted statistical values and obtains the statistical values.

For example, when receiving the encrypted statistical value "Enc (501)+Enc (530)", the data source A performs the homomorphic decryption $DEC (Enc (501)+Enc (530)) = 1031$. This 1031 is a statistical value of the sum of vehicle insurance scores of female users whose number of
10 claims is greater than 5.

In the data statistics method according to this example, the homomorphically encrypted data is sent to the cooperative data party during the data statistics, so that the cooperative data party can perform the statistical processing on the encrypted data. The statistical values can be obtained as long as the local homomorphic decryption is performed. This method protects the
15 privacy security of the local data and enables processing of the combined data of two data sources without exposing any data privacy. It also protects privacy security of all filtering fields by using a key exchange protocol.

In addition, the process as shown in the preceding FIG. 2 takes an example of a statistical value being the sum of a plurality of pieces of first data. In other examples, the data statistics
20 method according to one or more embodiments of the description is also applied to other statistical calculation scenarios. For example, a statistical value is an average value of a plurality of pieces of first data.

The process as shown in FIG. 2 can also be used for obtaining an "average of vehicle insurance scores of female users whose number of claims is greater than 5". The difference is
25 that, when performing the statistical processing on the various encrypted data, the data source B gets an average of the various encrypted data. For example, the value of Enc (501)+Enc (530) is divided by 2 to obtain an encrypted statistical value. After receiving the encrypted statistical value returned by the data source B, the data source A performs the homomorphic decryption to obtain the average.

30 To implement the preceding method, one or more embodiments of the description also provide a data statistics apparatus, and the apparatus is used for the data statistics between a local

data party and a cooperative data party. The local data party has the first data corresponding to the data identifiers, and the cooperative data party has the second data corresponding to the same data identifiers. In addition, the apparatus is configured to obtain the statistical values from a plurality of pieces of first data. As shown in FIG. 3, this apparatus comprises a data encryption module 31, a data sending module 32, a data receiving module 33, and a decryption processing module 34.

The data encryption module 31 is used to respectively perform the homomorphic encryption on the plurality of pieces of first data, and obtain the encrypted data.

The data sending module 32 is used to send the data identifiers corresponding to the plurality of pieces of first data and the encrypted data to the cooperative data party.

The data receiving module 33 is used to receive the encrypted statistical values returned by the cooperative data party. The encrypted statistical values are obtained by the cooperative data party by performing the statistical calculation on the encrypted data corresponding to the various data identifiers in an identifier intersection, wherein the identifier intersection includes the identifiers corresponding to the pieces of second data of the cooperative data party for the data statistics and selected from the plurality of data identifiers corresponding to the plurality of pieces of first data.

The decryption processing module 34 is used to perform the homomorphic decryption on the encrypted statistical values and obtain the statistical values.

To implement the preceding method, one or more embodiments of the description also provide a data statistics apparatus, and the data statistics apparatus is used for the data statistics between a local data party and a statistical data party. The statistical data party has a plurality of pieces of first data for calculating the statistical values, and the plurality of pieces of first data respectively corresponds to different data identifiers. The local data party has a plurality of pieces of second data corresponding to the data identifiers. As shown in FIG. 4, this apparatus comprises a data obtaining module 41, an intersection determining module 42, a statistical processing module 43, and a statistical sending module 44.

The data obtaining module 41 is used to receive the data identifiers and the corresponding encrypted data from the statistical data party. The data identifiers are the identifiers corresponding to the plurality of pieces of first data of the statistical data party for the data

statistics, and the encrypted data is obtained by the statistical data party by performing the homomorphic encryption on the plurality of pieces of first data.

The intersection determining module 42 is used to determine the identifier intersection according to the data identifiers corresponding to the plurality of pieces of second data of the local data party for the data statistics as well as the data identifiers of the plurality of pieces of first data.

The statistical processing module 43 is used to perform the statistical processing on the encrypted data corresponding to the various data identifiers in the identifier intersection, and obtain the encrypted statistical values.

The statistical sending module 44 is used to send the encrypted statistical values to the statistical data party, so that the statistical data party can perform the homomorphic decryption on the encrypted statistical values and obtain the statistical values.

To implement the preceding method, one or more embodiments of the description also provide a data statistics apparatus as shown in FIG. 5. The apparatus may comprise a data encryption module 51, an identifier processing module 52, a data sending module 53, an identifier receiving module 54, a private key processing module 55, a statistical receiving module 56, and a data decryption module 57.

The data encryption module 51 is used to perform the homomorphic encryption on a plurality of pieces of first data of local data statistics, and obtain the encrypted data.

The identifier processing module 52 is used to perform the local private key processing on the data identifiers respectively corresponding to the plurality of pieces of first data according to a key exchange protocol, and obtain a plurality of first processing identifiers.

The data sending module 53 is used to send the plurality of first processing identifiers and the encrypted data to the cooperative data party, so that the cooperative data party can generate the first key processing identifiers and store the corresponding relations between the first key processing identifiers and the encrypted data after performing peer private key processing on first processing identifiers.

The identifier receiving module 54 is used to receive second processing identifiers from the cooperative data party. The second processing identifiers are obtained by the cooperative data party by performing the peer private key processing on the data identifiers of the second data of data statistics.

The private key processing module 55 is used to generate and send the second key processing identifiers to the cooperative data party after performing the local private key processing on the second processing identifiers.

5 The statistical receiving module 56 is used to receive encrypted statistical values from the cooperative data party. The encrypted statistical values can be obtained by the cooperative data party by performing statistical processing on the various encrypted data corresponding to an identifier intersection between the first and second key processing identifiers.

The data decryption module 57 is used to perform the homomorphic decryption on the encrypted statistical values and obtain the statistical values.

10 In one example, the apparatus further comprises: a data filtering module that is used to obtain a plurality of pieces of first data of the local data statistics according to the predetermined data filtering conditions.

To implement the preceding method, one or more embodiments of the description also provide a data statistics apparatus as shown in FIG. 6. The apparatus may comprise a data
15 receiving module 61, a relation storage module 62, a private key encryption module 63, an identifier sending module 64, a statistical processing module 65, and a data sending module 66.

The data receiving module 61 is used to receive a plurality of peer processing identifiers and a plurality of pieces of encrypted data respectively corresponding to the plurality of peer processing identifiers from the statistical data party. The peer processing identifiers can be
20 obtained by the statistical data party by performing the peer private key processing on the data identifiers of the first data of the data statistics according to a key exchange protocol, and the encrypted data can be obtained by the statistical data party by performing the homomorphic encryption on the first data.

The relation storage module 62 is used to generate the first key processing identifiers and
25 store the corresponding relations between the first key processing identifiers and the corresponding encrypted data after performing the local private key operations on the various first processing identifiers according to a key exchange protocol.

The private key encryption module 63 is used to perform the local private key processing on the data identifiers respectively corresponding to the plurality of pieces of second data of the
30 local data statistics according to the key exchange protocol, and obtain a plurality of second processing identifiers.

The identifier sending module 64 is used to send the second processing identifiers to the statistical data party, and receive second key processing identifiers returned by the statistical data party. The second key processing identifiers can be obtained by the statistical data party by performing the peer private key processing on the second processing identifiers.

5 The statistical processing module 65 is used to obtain the various encrypted data corresponding to an intersection between the first and the second key processing identifiers, perform the statistical processing on the encrypted data, and obtain encrypted statistical values.

10 The data sending module 66 is used to send the encrypted statistical values to the statistical data party, so that the statistical data party can perform the homomorphic decryption on the encrypted statistical values and obtain the statistical values.

When the preceding apparatuses are described, various modules are respectively described in terms of functions for the convenience of the description. Of course, the functions of the various modules can be implemented in one or more pieces of software and/or hardware when one or more embodiments of the description are implemented.

15 For various steps in the processes as shown in the embodiments of the preceding method, their execution sequences are not limited to the sequences in the flow charts. In addition, the description of the various steps is implemented in a form of software, hardware, or a combination thereof. For example, technicians in the field implement it in a form of software codes that implements executable computer instructions of logic functions corresponding to the steps. When implemented in the form of software, the executable instructions are stored in the memory and executed by the processor in the apparatus.

20 For example, corresponding to the preceding method, one or more embodiments of the description provide a data statistics apparatus at the same time. The apparatus may comprise a processor, a memory, and computer instructions stored on the memory and running on the processor. The processor may be used to implement the following steps by executing the instructions:

respectively performing homomorphic encryption on the plurality of pieces of first data to obtain encrypted data;

25 sending the data identifiers corresponding to the plurality of pieces of first data and the encrypted data to the cooperative data party;

receiving encrypted statistical values returned by the cooperative data party, the encrypted statistical values being obtained by the cooperative data party by performing a statistical calculation on encrypted data corresponding to various data identifiers in an identifier intersection, the identifier intersection being identifiers corresponding to pieces of second data of the cooperative data party for the data statistics and selected from the plurality of data identifiers corresponding to the plurality of pieces of first data; and

performing homomorphic decryption on the encrypted statistical values and obtaining the statistical values.

The apparatus or modules described in the preceding embodiments are specifically implemented by a computer chip or entity, or by a product with a function. A computer is a typical implementation apparatus. The specific types of the computer include PCs, laptop computers, cell phones, camera phones, smart phones, personal digital assistants, media players, navigation apparatuses, email receiving and sending apparatuses, game consoles, tablet computers, wearable apparatuses, or the combination of any of these apparatuses.

Technicians in this field should understand that one or more embodiments of the description are provided as a method, a system, or a computer program product. Therefore, one or more embodiments of the description may use the form of hardware-only embodiments, software-only embodiments, or software and hardware combined embodiments. In addition, one or more embodiments of the description may use the form of computer program products implemented on one or more storage media (including but not limited to disk memories, CD-ROMs, and optical memories) containing computer program codes.

These computer program instructions may be also stored in a computer readable memory that can instruct computers or other programmable data processing apparatuses to operate in a particular way, so that instructions stored in the computer readable memory can generate products including an instruction device. The instruction device implements any specific functions in one or more processes in the flow charts and/or in one or more blocks in the block diagrams.

These computer program instructions may be also loaded to a computer or other programmable data processing equipment, so that a series of operating steps can be executed on the computer or other programmable equipment to generate processes implemented by the computer. Accordingly, the instructions executed in the computer or other programmable

equipment provide steps to implement any specific functions in one or more processes in the flow charts and/or in one or more blocks in the block diagrams.

It is important to note that the terms "comprise", "include" or any other variants of them are intended to cover a non-exclusive inclusion so that a process, a method, a commodity, or an apparatus comprising a series of elements comprises those elements and other elements that are not explicitly listed, or inherent elements of this process, method, commodity, or apparatus. Without further restrictions, elements restricted by the sentence "comprise a..." do not exclude the existence of other equal elements comprised in the process, method, commodity, or apparatus.

One or more embodiments of the description may be described in the general context of executable computer instructions executed by a computer, such as a program module. A program module generally comprises routines, programs, objects, components, and data structures that perform specific tasks or implement specific types of abstract data. One or more embodiments of the description may be also implemented in distributed computing environments where the tasks are performed by remote processing equipment connected through a communication network. In the distributed computing environments, the program modules may be located in local and remote computer storage media including memory devices.

Various embodiments of the description are described in a progressive manner, and the same and similar parts between the embodiments may be understood by cross references. The differences of an embodiment from other embodiments are emphasized in the description. In particular, the embodiments of the server-side apparatus are briefly described, because they are basically similar to the embodiments of the method. The relevant points may be referred to the description of the embodiments of the method.

Specific embodiments of this description are described above. Other embodiments are within the scope of the attached claims. In some cases, the operations or steps recorded in the claims may be performed in a sequence different from the sequence in the embodiments and still achieve the expected results. In addition, the processes depicted in the figures do not necessarily require a specific or sequential sequence shown in the figures to achieve the expected results. In some embodiments, multitasking and parallel processing are also possible or may be beneficial.

The foregoing is only one or more preferred embodiments of the description and is not intended to limit this invention. Any modification, equivalent substitution, improvement, etc.,

made within the spirit and principles of the present invention shall be included in the protection scope of the present invention.

Claims

1. A data statistics method, the method being applied to data statistics by combining data of a local data party and a cooperative data party, the local data party having a plurality of pieces of first data for calculating statistical values, the plurality of pieces of first data
5 respectively corresponding to different data identifiers, and the cooperative data party having a plurality of pieces of second data corresponding to the data identifiers, the method comprising:

respectively performing homomorphic encryption on the plurality of pieces of first data to obtain encrypted data;

10 sending the data identifiers corresponding to the plurality of pieces of first data and the encrypted data to the cooperative data party;

receiving encrypted statistical values returned by the cooperative data party, the encrypted statistical values being obtained by the cooperative data party by performing a statistical calculation on encrypted data corresponding to various data identifiers in an
15 identifier intersection, the identifier intersection being identifiers corresponding to pieces of second data of the cooperative data party for the data statistics and selected from the plurality of data identifiers corresponding to the plurality of pieces of first data; and

performing homomorphic decryption on the encrypted statistical values and
20 obtaining the statistical values.

2. The method according to claim 1,

wherein the sending data identifiers respectively corresponding to the plurality of pieces of first data to the cooperative data party comprises:

25 generating local private keys according to a key exchange protocol; and

performing local private key processing on the data identifiers using the local private keys to generate processed data identifiers, and sending the processed data identifiers to the cooperative data party; and

wherein the method further comprises:

30 receiving data identifiers sent from the cooperative data party and processed by private key processing of the cooperative data party; and

performing the local private key processing on the data identifiers received from the cooperative data party, and sending the data identifiers processed by the local private key processing to the cooperative data party.

35

3. A data statistics method, the method being applied to data statistics by combining data of a local data party and a statistical data party, the statistical data party having a plurality of pieces of first data for calculating statistical values, the plurality of pieces of first data respectively corresponding to different data identifiers, and the local data party having a plurality of pieces of second data corresponding to the data identifiers, the method comprising:

receiving data identifiers and corresponding encrypted data from the statistical data party, the data identifiers being identifiers corresponding to pieces of first data of the statistical data party for the data statistics, and the encrypted data being obtained by the statistical data party by performing homomorphic encryption on the pieces of first data;

determining an identifier intersection according to data identifiers corresponding to pieces of second data of the local data party for the data statistics and the data identifiers of the pieces of first data;

performing statistical processing on encrypted data corresponding to various data identifiers in the identifier intersection to obtain encrypted statistical values; and

sending the encrypted statistical values to the statistical data party to enable the statistical data party to perform homomorphic decryption on the encrypted statistical values to obtain the statistical values.

4. The method according to claim 3, wherein the data identifiers received from the statistical data party are data identifiers generated by the statistical data party performing private key processing of the statistical data party on the data identifiers corresponding to the pieces of first data;

wherein after receiving the data identifiers from the statistical data party, the method further comprises:

generating local private keys according to a key exchange protocol;

generating first key processing identifiers by performing local private key processing using the local private keys on the data identifiers processed by the private key processing of the statistical data party; and

storing corresponding relations between the first key processing identifiers and the encrypted data;

wherein the method further comprising:

performing the local private key processing on data identifiers corresponding to the pieces of second data of the local data party for the data statistics using the local private keys, and send the processed data identifiers to the statistical data party; and

receiving second key processing identifiers returned by the statistical data party,
wherein the second key processing identifiers are generated by the private key processing
of the statistical data party; and

wherein the determining an identifier intersection according to data identifiers

- 5 corresponding to pieces of second data of the local data party for the data statistics and the
data identifiers of the pieces of first data, comprises: determining an identifier intersection
between the first key processing identifiers and the second key processing identifiers.

- 10 5. A data statistics method, the method being used for data statistics between a local data
party and a cooperative data party, the local data party having first data corresponding to data
identifiers, the cooperative data party storing second data corresponding to the same data
identifiers, and the method being applied to obtain statistical values from a plurality of pieces
of first data, the method comprising:

- 15 performing homomorphic encryption on the plurality of pieces of first data of the local
data party for the data statistics to obtain encrypted data;

performing local private key processing on data identifiers respectively corresponding to
the plurality of pieces of first data according to a key exchange protocol to obtain local
processing identifiers;

- 20 sending the local processing identifiers and the encrypted data corresponding to the
pieces of first data to the cooperative data party, to enable the cooperative data party to
generate first key processing identifiers by performing peer private key processing on the
local processing identifiers, and store corresponding relations between the first key processing
identifiers and the encrypted data;

- 25 receiving second processing identifiers from the cooperative data party, the second
processing identifiers being obtained by the cooperative data party by performing the peer
private key processing on data identifiers corresponding to second data for the data statistics;

performing the local private key processing on the second processing identifiers to
generate second key processing identifiers, and sending the second key processing identifiers
to the cooperative data party;

- 30 receiving encrypted statistical values from the cooperative data party, the encrypted
statistical values being obtained by the cooperative data party by performing statistical
processing on various encrypted data corresponding to an identifier intersection between the
first key processing identifiers and the second key processing identifiers; and

- 35 performing homomorphic decryption on the encrypted statistical values and obtaining
the statistical values.

6. The method according to claim 5, wherein
the plurality of pieces of first data is located in a same statistical column of the local data
party.

5

7. The method according to claim 5, further comprising:
obtaining the plurality of pieces of first data of local data statistics according to
predetermined data filtering conditions.

10 8. A data statistics method, the method being used for data statistics between a local data
party and a statistical data party, the statistical data party having first data corresponding to
data identifiers, the local data party storing second data corresponding to the data identifiers,
and the method being applied to obtain statistical values from a plurality of pieces of first
data, the method comprising:

15 receiving a plurality of peer processing identifiers and a plurality of pieces of encrypted
data respectively corresponding to the plurality of peer processing identifiers from the
statistical data party, the plurality of peer processing identifiers being obtained by the
statistical data party by performing peer private key processing on data identifiers of the
plurality of pieces of first data for the data statistics according to a key exchange protocol, the
20 plurality of pieces of encrypted data being obtained by the statistical data party by performing
homomorphic encryption on the plurality of pieces of first data;

performing local private key processing on the plurality of peer processing identifiers
according to a key exchange protocol for generating first key processing identifiers, and
storing corresponding relations between the first key processing identifiers and the
25 corresponding encrypted data;

performing the local private key processing on data identifiers respectively
corresponding to a plurality of pieces of second data of the local data party for the data
statistics according to the key exchange protocol to obtain second processing identifiers;
sending the second processing identifiers to the statistical data party, and receiving
30 second key processing identifiers returned by the statistical data party, the second key
processing identifiers being obtained by the statistical data party by performing the peer
private key processing on the second processing identifiers;

obtaining various encrypted data corresponding to an identifier intersection between the
first key processing identifiers and the second key processing identifiers, and performing
35 statistical processing on the various encrypted data to obtain encrypted statistical values; and

sending the encrypted statistical values to the statistical data party for the statistical data party to perform homomorphic decryption on the encrypted statistical values to obtain the statistical values.

5 9. The method according to claim 8, wherein when the statistical values are sums of the plurality of pieces of first data, the performing statistical processing on the various encrypted data to obtain encrypted statistical values comprises:

performing a summation statistical calculation on the various encrypted data.

10 10. The method according to claim 8, wherein when the statistical values are average values of the plurality of pieces of first data, the performing statistical processing on the various encrypted data to obtain encrypted statistical values comprises:

performing an average value calculation on the various encrypted data.

15 11. A data statistics apparatus, the apparatus being used for data statistics between a local data party and a cooperative data party, the local data party having first data corresponding to data identifiers, the cooperative data party having second data corresponding to the data identifiers, and the apparatus being applied to obtain statistical values from a plurality of pieces of first data, the apparatus comprising:

20 a data encryption module that is used to respectively perform homomorphic encryption on the plurality of pieces of first data to obtain encrypted data;

a data sending module that is used to send data identifiers corresponding to the plurality of pieces of first data and the encrypted data to the cooperative data party;

25 a data receiving module that is used to receive encrypted statistical values returned by the cooperative data party, the encrypted statistical values being obtained by the cooperative data party by performing a statistical calculation on encrypted data corresponding to various data identifiers in an identifier intersection, the identifier intersection being identifiers corresponding to pieces of second data of the cooperative data party for the data statistics and selected from the plurality of data identifiers corresponding to the plurality of pieces of first data; and

30 a decryption processing module that is used to perform homomorphic decryption on the encrypted statistical values and obtain the statistical values.

35 12. A data statistics apparatus, the apparatus being used for data statistics between a local data party and a statistical data party, the statistical data party having a plurality of pieces of

first data for calculating statistical values, the plurality of pieces of first data respectively corresponding to different data identifiers, and the local data party having a plurality of pieces of second data corresponding to the data identifiers, the apparatus comprising:

- 5 a data obtaining module that is used to receive data identifiers and corresponding encrypted data from the statistical data party, the data identifiers being identifiers corresponding to pieces of first data of the statistical data party for the data statistics, and the encrypted data being obtained by the statistical data party by performing homomorphic encryption on the pieces of first data;
- 10 an intersection determining module that is used to determine an identifier intersection according to data identifiers corresponding to pieces of second data of the local data party for the data statistics and the data identifiers of the pieces of first data;
- a statistical processing module that is used to perform statistical processing on encrypted data corresponding to various data identifiers in the identifier intersection to obtain encrypted statistical values; and
- 15 a statistical sending module that is used to send the encrypted statistical values to the statistical data party to enable the statistical data party to perform homomorphic decryption on the encrypted statistical values to obtain the statistical values.

- 20 13. A data statistics apparatus, the apparatus being used for data statistics between a local data party and a cooperative data party, the local data party having first data corresponding to data identifiers, the cooperative data party storing second data corresponding to the data identifiers, and the apparatus being applied to obtain statistical values from a plurality of pieces of first data, the apparatus comprising:

- 25 a data encryption module that is used to perform homomorphic encryption on the plurality of pieces of first data of the local data party for the data statistics to obtain encrypted data;
- an identifier processing module that is used to perform local private key processing on data identifiers respectively corresponding to the plurality of pieces of first data according to a key exchange protocol to obtain local processing identifiers;
- 30 a data sending module that is used to send the local processing identifiers and the encrypted data to the cooperative data party for the cooperative data party to generate first key processing identifiers by performing peer private key processing on the local processing identifiers, and store corresponding relations between the first key processing identifiers and the encrypted data;

an identifier receiving module that is used to receive second processing identifiers from the cooperative data party, the second processing identifiers being obtained by the cooperative data party performing the peer private key processing on data identifiers of a plurality of pieces of second data for the data statistics;

5 a private key processing module that is used to perform the local private key processing on the second processing identifiers to generate second key processing identifiers, and send the second key processing identifiers to the cooperative data party;

a statistical receiving module that is used to receive encrypted statistical values from the cooperative data party, the encrypted statistical values being obtained by the cooperative data party by performing statistical processing on various encrypted data corresponding to an intersection between the first key processing identifiers and the second key processing identifiers; and

a data decryption module that is used to perform homomorphic decryption on the encrypted statistical values and obtain statistical values.

15

14. A data statistics apparatus, the apparatus being used for data statistics between a local data party and a statistical data party, the statistical data party storing first data corresponding to data identifiers, the local data party storing second data corresponding to the data identifiers, and the apparatus being applied to obtain statistical values from a plurality of pieces of first data, the apparatus comprising:

20 a data receiving module that is used to receive a plurality of peer processing identifiers and a plurality of pieces of encrypted data respectively corresponding to the plurality of peer processing identifiers from the statistical data party, the plurality of peer processing identifiers being obtained by the statistical data party by performing peer private key processing on data identifiers of the plurality of pieces of first data for the data statistics according to a key exchange protocol, the plurality of pieces of encrypted data being obtained by the statistical data party by performing homomorphic encryption on the first data;

25 a relation storage module that is used to perform local private key processing on the plurality of peer processing identifiers according to a key exchange protocol to generate first key processing identifiers, and store corresponding relations between the first key processing identifiers and the corresponding encrypted data;

30 a private key encryption module that is used to perform the local private key processing on data identifiers respectively corresponding to a plurality of pieces of second data of the local data party for the data statistics according to the key exchange protocol to obtain second processing identifiers;

35

an identifier sending module that is used to send the second processing identifiers to the statistical data party, and receive second key processing identifiers returned by the statistical data party, the second key processing identifiers being obtained by the statistical data party by performing the peer private key processing on the second processing identifiers;

5 a statistical processing module that is used to obtain various encrypted data corresponding to an identifier intersection between the first key processing identifiers and the second key processing identifiers, and perform statistical processing on the various encrypted data to obtain encrypted statistical values; and

10 a data sending module that is used to send the encrypted statistical values to the statistical data party for the statistical data party to perform homomorphic decryption on the encrypted statistical values and obtain the statistical values.

15 15. A data statistics apparatus, the said apparatus comprising a memory, a processor, and computer instructions stored on the memory and running on the processor, the processor implementing the steps during the execution of instructions including:

 respectively performing homomorphic encryption on the plurality of pieces of first data to obtain encrypted data;

 sending data identifiers corresponding to the plurality of pieces of first data and the encrypted data to a cooperative data party;

20 receiving encrypted statistical values returned by the cooperative data party, the encrypted statistical values being obtained by the cooperative data party by performing a statistical calculation on encrypted data corresponding to various data identifiers in an identifier intersection, the identifier intersection being identifiers corresponding to pieces of second data of the cooperative data party for data statistics and selected from the plurality of

25 data identifiers corresponding to the plurality of pieces of first data; and

 performing homomorphic decryption on the encrypted statistical values and obtaining statistical values.