(12)Indian Patent Application

---

(21) Application Number: 6206/CHE/2014

(22) Filing Date: 09/12/2014 (43) Publication Date: 26/08/2016

(71) Applicant(s): L & T TECHNOLOGHY SERVICES LIMITED

(72) Inventor(s): DIVYA, NAIR G

(51) International Classifications: H04W

(54) Title: SYSTEM AND METHOD FOR SECURELY TRANSMITTING ANONYMOUS INFORMATION

(57) Abstract: A system and method for securely transmitting anonymous information for a plurality of items over a communication network is disclosed. The disclosed system and method receives data from a plurality of users for the plurality of items through one or more user interfaces. The data received from each of the users is encoded into a predefined binary format. A decimal equivalent is identified corresponding to each of the binary encoded data. A secret sharing scheme is applied to each of the decimal encoded data to generate a predefined number of shares. The secret sharing scheme uses a polynomial construction for generating the shares. Each share of the decimal encoded data is sent to a different repository through the communication network. The number of repositories is equal to the number of shares generated for one decimal encoded data. The shares are summed up at each of the repositories. The summation value from two or more repositories are selected based on a predefined threshold scheme and the selected summation values are interpolated to obtain a polynomial of the encoded data. A constant term in the polynomial is decoded to reconstruct transmitted information.

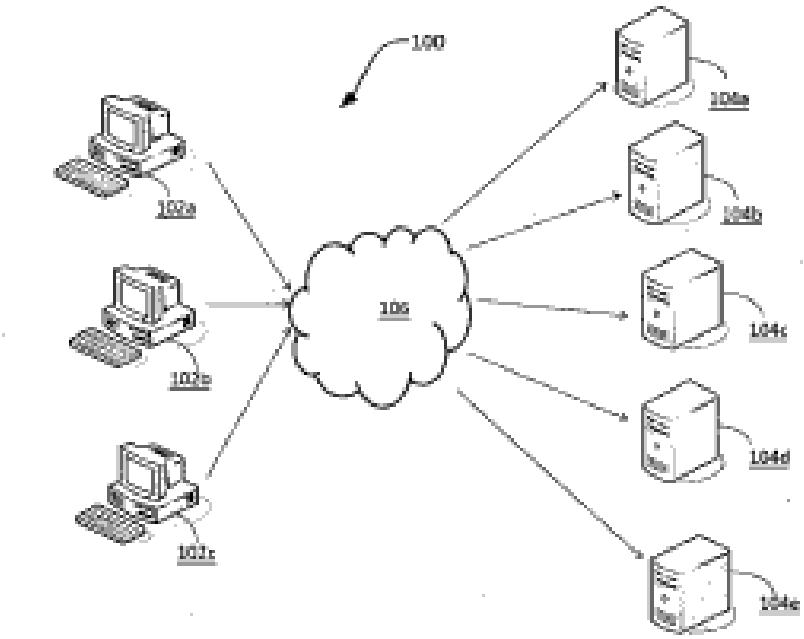


Figure 1

ABSTRACT

A system and method for securely transmitting anonymous information for a plurality of items over a communication network is disclosed. The disclosed system and method receives data from a plurality of users for the plurality of items through one or more user interfaces. The data received from each of the users is encoded into a predefined binary format. A decimal equivalent is identified corresponding to each of the binary encoded data. A secret sharing scheme is applied to each of the decimal encoded data to generate a predefined number of shares. The secret sharing scheme uses a polynomial construction for generating the shares. Each share of the decimal encoded data is sent to a different repository through the communication network. The number of repositories is equal to the number of shares generated for one decimal encoded data. The shares are summed up at each of the repositories. The summation value from two or more repositories are selected based on a predefined threshold scheme and the selected summation values are interpolated to obtain a polynomial of the encoded data. A constant term in the polynomial is decoded to reconstruct transmitted information.

Claims

We claim

1. A method for securely transmitting anonymous information for a plurality of items over a communication network, the method comprising:

receiving data from a plurality of users for the one or more items through a user interface;

encoding the data received from each of the users into a predefined binary format;

identifying a corresponding decimal equivalent of each of the binary encoded data;

applying a secret sharing scheme to the each of the decimal encoded data to generate a predefined number of shares, the secret sharing scheme using a polynomial construction;

sending each share of the decimal encoded data to a different repository through the communication network, number of repositories being equal to the number of shares generated for one decimal encoded data;

summing up the shares received at each of the repositories;

selecting the summation values from two or more repositories based on a predefined threshold and interpolating the selected summation values to obtain a polynomial of the encoded data;

decoding a constant term in the polynomial to reconstruct transmitted information.

2. The method of claim1 wherein the secret sharing scheme is based on Shamir's or Blakley's scheme.

3. The method of claim1 wherein encoding the data includes identifying the size of the binary pattern according to a formula.

4. The method of claim1 wherein the communication network is a wireless network.

5. The method of claim1 wherein the interpolating the selected summation values is based on Lagrange interpolation.

6. The method of claim1 wherein the predefined binary format includes a format that is directly proportional to number of items and users.

7. The method of claim1 wherein the user interface is a graphical user interface.

8. The method of claim1 wherein the repository is a collection centre or warehouse.

9. The method of claim1 wherein the predefined threshold is based on Shamir's (m, k) threshold scheme.

10. The method of claim 5 wherein the Lagrange interpolation formula used is

$$q(x) = \sum_{i=1}^{k} y_i \prod_{j=1, j=i}^{k} \frac{x - x_j}{x_i - x_j}$$

where q(x) is the polynomial obtained by the interpolation of the summation values from the selected two or more repositories, x_i corresponds to i^{th} share of the secret and y_i corresponds to value of the i^{th} share and i and j are two variables which ranges from 1 to k where k is number of shares required to reconstruct the secret.

Dated this 30th day of November 2015

Mohammed Faisal (INPA No: 1941)
Head, IPR
L&T Technology Services Limited

FIELD OF INVENTION

The present invention relates in general to a system and method for securing data from maliciously caused destruction, unauthorized modification, or unauthorized disclosure.

BACKGROUND

One way to secure data from unauthorized access or unauthorized use is to use a secret sharing scheme. A secret sharing scheme is a method to split a sensitive piece of data (e.g., confidential files, an encryption key, or any type of communication), sometimes called the secret, into a collection of pieces, called shares, such that possession of a sufficient number of shares enables recovery of the secret, but possession of an insufficient number of shares provides little or no information about the secret that was shared. Such schemes are important tools in cryptography and information security.

Secure multi-party computation (SMC) is an active research area in cryptography. The SMC allows a set of parties to compute a function of their inputs while preserving input privacy and correctness. There may be cases when mutually distrustful parties may need to perform a joint computation but may not afford to reveal their inputs to each other. This may occur, for example, during auctions, data mining, voting, negotiations and business analytics.

In secret sharing, the secret is not single handed, but multi-handed so that even if any of the parties involved in the computation are malicious, the secret may be reconstructed. Development of secret sharing scheme started as a solution to the problem of safeguarding cryptographic keys by distributing the key among many participants, say 'n' and 't' or more of

the 'n' participants may recover it by pooling their shares. Thus the authorized set is any subset of participants containing more than 't' members. The scheme is denoted as (t, n) threshold scheme.

Existing security techniques are based on cryptographic algorithms which are computationally complex. The existing approaches use encryption keys or homographic functions for cryptography.

SUMMARY OF THE INVENTION

The invention provides a new mechanism for anonymization that is grounded on secret sharing based secure multi-party computation (SMC).

According to an embodiment of the invention, a system and method for securely transmitting anonymous information for a plurality of items over a communication network is disclosed. According to an exemplary embodiment, the disclosed system and method receives data from a plurality of users for the plurality of items through one or more user interfaces. The data received from each of the users is encoded into a predefined binary format. A decimal equivalent is identified corresponding to each of the binary encoded data. A secret sharing scheme is applied to each of the decimal encoded data to generate a predefined number of shares. The secret sharing scheme uses a polynomial construction for generating the shares. Each share of the decimal encoded data is sent to a different repository through the communication network. The number of repositories is equal to the number of shares generated for one decimal encoded data. The shares are summed up at each of the repositories. The summation value from two or more repositories are selected based on a predefined threshold

scheme and the selected summation values are interpolated to obtain a polynomial of the encoded data. A constant term in the polynomial is decoded to reconstruct transmitted information.

BRIEF DESCRIPTION OF DRAWINGS

Other objects, features, and advantages of the invention will be apparent from the following description when read with reference to the accompanying drawings. In the drawings, wherein like reference numerals denote corresponding parts throughout the several views:

Figure 1 illustrates an exemplary system architecture for securely transmitting anonymous information for a plurality of items over a communication network according to an embodiment of the invention;

Figure 2 illustrates an exemplary feedback system architecture for review of a set of products by a set of reviewers according to an embodiment of the invention; and

Figure 3 illustrates an exemplary process flow for securely transmitting anonymous information for a plurality of items over a communication network in accordance with one embodiment of the invention.

DETAILED DESCRIPTION OF DRAWINGS

According to embodiments of the invention, a system and method for securely transmitting anonymous information is disclosed. The disclosed system and method are grounded on secret sharing based secure multi-party computation (SMC).

According to an embodiment of the invention, Figure 1 illustrates a system 100 for securely transmitting anonymous information. As illustrated, the system 100 may have one or more user interface 102a, 102b, 102c for receiving data from a plurality of users. The data may be pertinent to anonymous information about one or more items from a plurality of items. According to an embodiment, the items may correspond to any person, product or services or a combination thereof. According to another embodiment, the data may correspond to feedback regarding a particular product or service and the user may be a voter or reviewer.

The data received from each of the users may be encoded into a predefined binary format. According to an embodiment, the predefined binary format may include a format that may be directly proportional to the number of products and reviewers.

Exemplary, consider a system for review of a product by selecting a particular grade from a set of grades including Excellent (E), Good (G), Average (A) and Poor (P) by a set of say five reviewers, say R1, R2, R3, R4 and R5. The result of the review process may be total number of points in each category without knowing who selected which. The process starts with identifying the predefined binary format in which the data received from each reviewer may be encoded. To encode the data, size of bit pattern may be identified. The size of bit pattern may be set according to $n(1+\log_2 m)$ and product bit block size may be set according to $(1+\log_2 m)$ where m corresponds to number of reviewers, n corresponds to number of grades and $\log_2 m$ corresponds to number of bits required to represent m in binary. The bit pattern may be of size $4*(1+\log_2 5)$ i.e. $4*(1+3)$ equal to sixteen bits. Table 1 illustrates the exemplary binary pattern after encoding.

Table 1

Reviewers	Excellent	Good	Average	Poor	After Encoding
R1	E				0001 0000 0000 0000

R2		G			0000 0001 0000 0000
R3		G			0000 0001 0000 0000
R4			A		0000 0000 0001 0000
R5	E				0001 0000 0000 0000

According to an embodiment, the data may be encoded by a processor. According to yet another embodiment, the processor may be a part of the user interface 102a, 102b, 102c.

After encoding data in binary digits, a corresponding decimal equivalent of the binary encoded data may be identified. Referring back to the aforementioned example, the decimal equivalent for binary pattern 0001 0000 0000 0000 is 4096; the decimal equivalent for binary pattern 0000 0001 0000 0000 is 256 and the decimal equivalent for binary pattern 0000 0000 0001 0000 is sixteen. Table 2 illustrates the corresponding decimal equivalents of the binary patterns discussed in Table 1.

Table 2

Reviewers	Excellent	Good	Average	Poor	After Encoding	Decimal Equivalent
R1	E				0001 0000 0000 0000	4096
R2		G			0000 0001 0000 0000	256
R3		G			0000 0001 0000 0000	256
R4			A		0000 0000 0001 0000	16
R5	E				0001 0000 0000 0000	4096

A secret sharing scheme may be applied to the decimal encoded data to generate a predefined number of shares. According to an embodiment, the secret sharing scheme may use a polynomial construction for generating the shares. According to yet another embodiment, the secret sharing may be based on Shamir's or Blakley's scheme.

Each share of the decimal encoded data may be sent to a predefined repository 104a, 104b, 104c, 104d, 104e through a communication network 106. The communication network may

be a wired network or a wireless communication network. According to an embodiment, the repository may be a collection centre or warehouse. The total number of repository may be proportional to the number of shares generated. Referring back to the aforementioned example, considering five shares are generated for each decimal encoded data then all the first share of each of the decimal encoded data may be sent to first repository 104a, all the second share may be sent to second repository 104b etc. Table 3 illustrates the decimal equivalent of the secret divided into shares and sent to five repositories CC1 to CC5.

Table 3

Secret	CC1 (104a)	CC2(104b)	CC3(104c)	CC4(104d)	CC5(104e)
4096	(1,4151)	(2,4254)	(3,4405)	(4,4604)	(5,4851)
256	(1,379)	(2,572)	(3,835)	(4,1168)	(5,1571)
256	(1,313)	(2,412)	(3,553)	(4,736)	(5,961)
16	(1,108)	(2,240)	(3,412)	(4,624)	(5,876)
4096	(1,4185)	(2,4320)	(3,4501)	(4,4728)	(5,5001)

The shares are summed up at each of the repositories 104a, 104b, 104c, 104d, 104e. A computation agent may sum up the shares in respective repositories 104a, 104b, 104c, 104d, 104e. Table 4 illustrates the summation at each of the repositories CC1 to CC5.

Table 4

Secret	CC1	CC2	CC3	CC4	CC5
4096	(1,4151)	(2,4254)	(3,4405)	(4,4604)	(5,4851)
256	(1,379)	(2,572)	(3,835)	(4,1168)	(5,1571)
256	(1,313)	(2,412)	(3,553)	(4,736)	(5,961)
16	(1,108)	(2,240)	(3,412)	(4,624)	(5,876)
4096	(1,4185)	(2,4320)	(3,4501)	(4,4728)	(5,5001)
	(1, 9136)	(2, 9798)	(3, 10706)	(4, 11860)	(5, 13260)

The two or more repositories 104a, 104b, 104c, 104d, 104e may be selected randomly based on a predefined threshold scheme. The threshold scheme may be (5, 3) threshold scheme. The summation values from the randomly selected repositories may be interpolated to obtain a polynomial of the encoded data. According to an embodiment, the interpolation formula may

be Lagrange interpolation. A constant term in the polynomial may be decoded to produce the result.

Say, for example, the repositories CC1, CC3 and CC5 may be randomly selected and sums of the shares received at CC1, CC3 and CC5 may be (1, 9136), (3, 10706) and (5, 13260) respectively.

The Lagrange interpolation formula may be used:

$$q(x) = \sum_{i=1}^{k} y_i \prod_{j=1, j=i}^{k} \frac{x - x_j}{x_i - x_j}$$

where q(x) is the polynomial obtained by the interpolation of the summation values from the randomly selected repositories, the random selection of repositories based on threshold k. x_i corresponds to i^{th} share of the secret and y_i corresponds to value of the i^{th} share. i and j are two variables which ranges from 1 to k where k is the number of shares required to reconstruct the secret.

Typically Lagrange formula results in a quadratic equation say AX^2+BX+C, where A, B, C are first, second and third constant respectively. To deduce result, we may utilize the third constant C. Suppose the shares selected for reconstruction are (x1, y1) = (1, 9136), (x2, y2) = (3, 10706) and (x3, y3) = (5, 13260). Using (x1, y1), (x2, y2) and (x3, y3) in the aforementioned Lagrange formula may result in polynomial $123\ x^2+293\ x+8720$.

Decoding constant 8720 may result in binary format 0010 0010 0001 0000 where first consecutive four bits refers to the reviews received by first item (Grade E), second consecutive four bits refers to the reviews received by second item (Grade G) and so on. Accordingly, the

reviews (E, G, A, P) received by the exemplary product may correspond to 2, 2, 1, 0 respectively.

According to another embodiment of the invention, Figure 2 illustrates a system 200 for securely transmitting anonymous information. As illustrated, the system 200 may have one or more user interface 202a, 202b, 202c, 202d for receiving data from a plurality of users. The data may be pertinent to anonymous information about one or more items from a plurality of items. According to an embodiment, the items may correspond to any person, product or services or a combination thereof. According to another embodiment, the data may correspond to feedback regarding a particular product or service and the user may be a voter or reviewer.

The system 200 may include a share generating module 208. The share generating module 208 receives data from the one or more user interface 202a, 202b, 202c, 202d and encodes the data into a predefined binary format. According to an embodiment, the predefined binary format may include a format that may be directly proportional to the number of products and reviewers.

Exemplary, consider the system for review of products where each product may be given a rating from every reviewer in a given range of values. The bit pattern may depend upon the range of rating and number of reviewers and products. Say for example, consider a system having five reviewers, say R1, R2, R3, R4 and R5 and three products, say P1, P2 and P3 and the range of rating may be in range of 1 to 4 where 1 is the least and 4 is the highest. Since the range of rating for the product is from 1 to 4, the maximum value one product may get is 4*5 = 20 and $\log_2 20$ is five, therefore, five bits may be required to represent each product. Table 5 illustrates the exemplary binary pattern after encoding.

Table 5

Reviewers	Product P1	Product P2	Product P3	After Encoding	Decimal Equivalent
R1	2	1	4	00100 00001 00010	4130
R2	1	3	2	00010 00011 00001	2145
R3	2	2	3	00011 00010 00010	3138
R4	1	2	1	00001 00010 00001	1089
R5	2	1	1	00001 00001 00010	1058

The five least significant bits represent rating by reviewers R1, R2, R3, R4 and R5 for product P1, middle five bits represent rating by reviewers R1, R2, R3, R4 and R5 for product P2 and five most significant bits represent rating by reviewers R1, R2, R3, R4 and R5 for product P3. For example, referring to first row of Table 5, '00010'represents rating for product P1 by reviewer R1, '00001' represents rating for product P2 by reviewer R1 and '00100'represents rating for product P3 by reviewer R1.

After encoding in binary digits, the share generating module 208 determines a corresponding decimal equivalent of the binary encoded data. Table 5 illustrates the corresponding decimal equivalents of the binary patterns. For example, the decimal equivalent corresponding to binary format 00100 00001 00010 is 4130.

The share generating module 208 applies a secret sharing scheme to the decimal encoded data illustrated in Table 5 to generate a predefined number of shares. According to an embodiment, the secret sharing scheme may use a polynomial construction for generating the shares. According to yet another embodiment, the secret sharing may be based on Shamir's or Blakley's scheme.

Each share of the decimal encoded data may be sent to a predefined repository 204a, 204b, 204c, 204d through a communication network 206. According to an embodiment, the repository may be a collection centre or warehouse. The total number of repository may be proportional to the number of shares generated.

The shares are summed up at each of the repositories 204a, 204b, 204c, 204d. A computation agent may sum up the shares in respective repositories 204a, 204b, 204c, 204d.

The two or more repositories 204a, 204b, 204c 204d may be selected randomly based on a predefined threshold scheme. For example, the threshold scheme may be (5, 3) threshold scheme. The summation values from the randomly selected repositories may be interpolated to obtain a polynomial of the encoded data. According to an embodiment, the interpolation formula may be Lagrange interpolation. A constant term in the polynomial may be decoded to produce the result.

Exemplary, say decoding the constant term in the polynomial may result in 01011 01001 01000 where the five least significant bits correspond to value 8 which may be the total points for Product P1 by reviewers R1, R2, R3, R4 and R5, next 5 bits may correspond to value 9 which may be the total points for Product P2 by reviewers R1, R2, R3, R4 and R5 and five most significant bits may correspond to value 11 which may be the total points for Product P3 by reviewers R1, R2, R3, R4 and R5.

One exemplary application of the invention may be in feedback management system which may be used for industrial organizations, educational institutions, government organizations

etc. to get the reviews or feedback regarding products or individuals. The other applications may be extended to many areas like E-voting, secure query retrieval and the like.

According to exemplary embodiments of the invention, the disclosed method and system does not try to hide the user; rather provides the secrecy of the user thereby securing the anonymity of the data. Through this mechanism, no one may get an idea regarding the particular feedback of the user, but only may get the final result.

FIG. 3 illustrates a block diagram of a cryptographic method 300 for securely transmitting anonymous information, according to an embodiment of the invention. The disclosed method includes receiving data from a plurality of users. The data may be pertinent to anonymous information about one or more items from a plurality of items.

At step 302, the data received from each of the users may be encoded in to a predefined binary format. According to an embodiment, the predefined binary format may include a format that may be directly proportional to the number of products and reviewers.

At step 304, a corresponding decimal equivalent of the binary encoded data may be identified.

At step 306, the decimal equivalent of the binary encoded data may be divided into a predefined number of shares based on a threshold. According to an embodiment, the secret sharing scheme may use a polynomial construction for generating the shares. According to yet another embodiment, the secret sharing may be based on Shamir's or Blakley's scheme.

At step 308, each share of the decimal encoded data may be sent to a predefined repository 104a, 104b, 104c, 104d, 104e through a communication network 106 as shown in Figure 1. According to an embodiment, the repository may be a collection centre or warehouse. The total number of repository may be proportional to the number of shares generated. Different shares may be sent through different communication links.

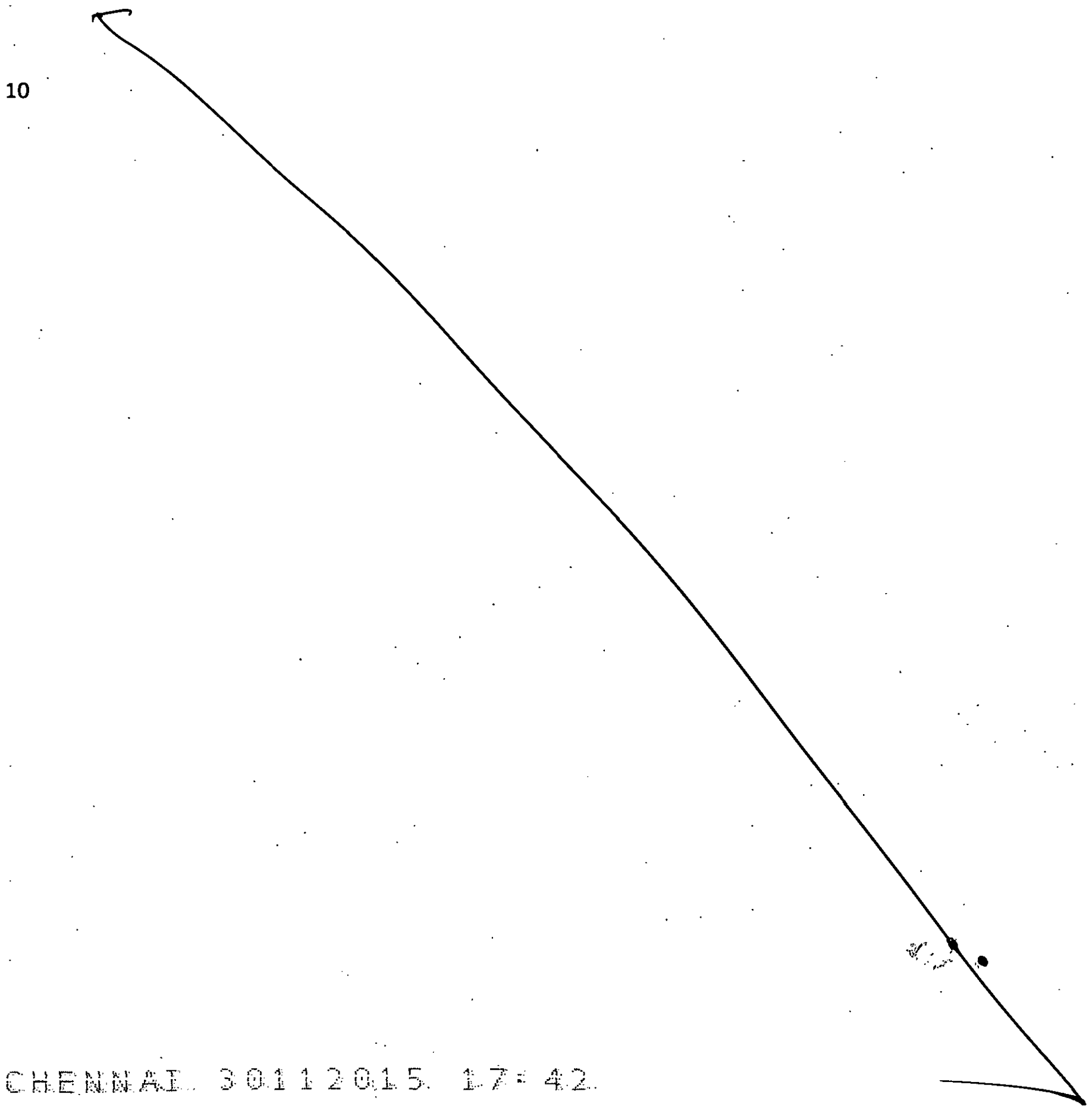
At step 310, shares may be summed up at each of the repositories 104a, 104b, 104c, 104d, 104e of Figure 1. A computation agent may sum up the shares in respective repositories 104a, 104b, 104c, 104d, 104e. The two or more repositories 104a, 104b, 104c, 104d, 104e may be selected randomly based on a predefined threshold scheme. The summation values from the randomly selected repositories may be interpolated to obtain a polynomial of the encoded data. According to an embodiment, the interpolation formula may be Lagrange interpolation. A constant term in the polynomial may be decoded to produce the result.

In the drawings and specification there has been set forth preferred embodiments of the invention, and although specific terms are employed, these are used in a generic and descriptive sense only and not for purposes of limitation. Changes in the form and the proportion of parts, as well as in the substitution of equivalents, are contemplated as circumstances may suggest or render expedient without departing from the spirit or scope of the invention.

Throughout the various contexts described in this disclosure, the embodiments of the invention further encompass computer apparatus, computing systems and machine-readable media configured to carry out the foregoing systems and methods. In addition to an embodiment consisting of specifically designed integrated circuits or other electronics, the present invention may be conveniently implemented using a conventional general purpose or a specialized digital

computer or microprocessor programmed according to the teachings of the present disclosure, as will be apparent to those skilled in the computer art.

Appropriate software coding can readily be prepared by skilled programmers based on the teachings of the present disclosure, as will be apparent to those skilled in the software art. The invention may also be implemented by the preparation of application specific integrated circuits or by interconnecting an appropriate network of conventional component circuits, as will be readily apparent to those skilled in the art.



700022408

L&T Technology Services Limited

6206/CHE/2014

Total number of Sheets: 3
Sheet No. 1 of 3

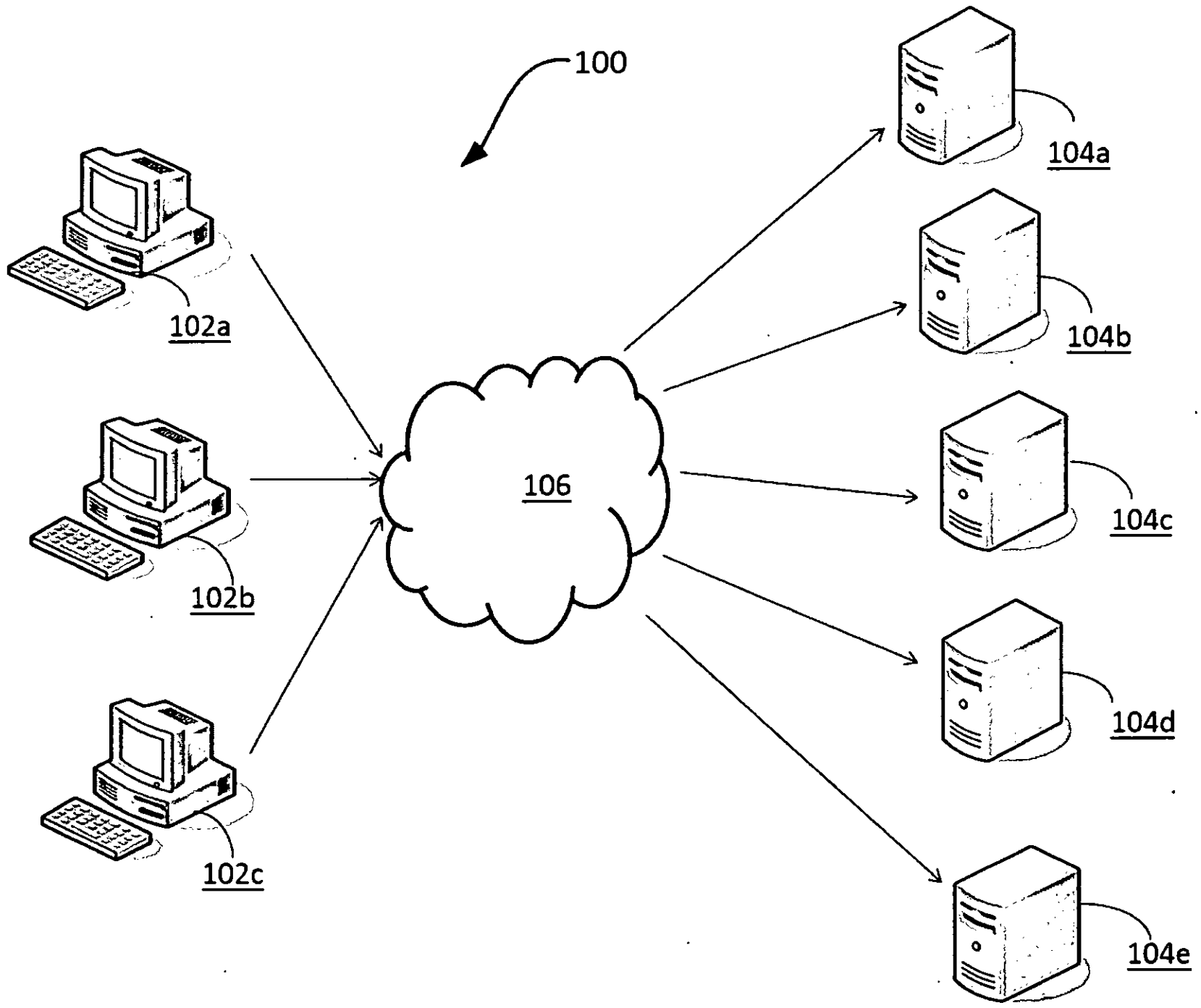


Figure 1

Mohammed Faisal (INPA No: 1941)
Head, IPR
L&T Technology Services Ltd.

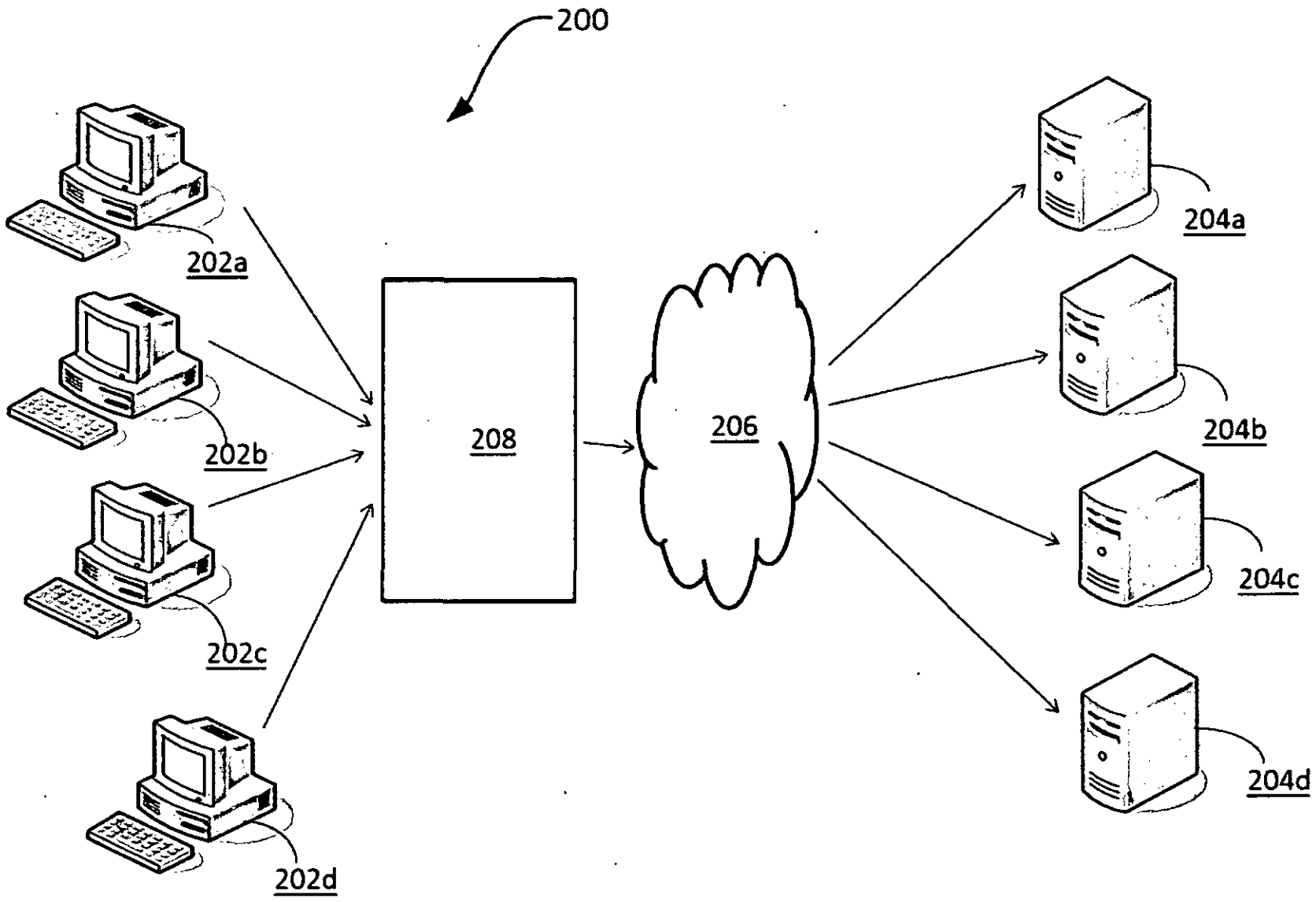


Figure 2

Mohammed Faisal (INPA No: 1941)
Head, IPR
L&T Technology Services Ltd.

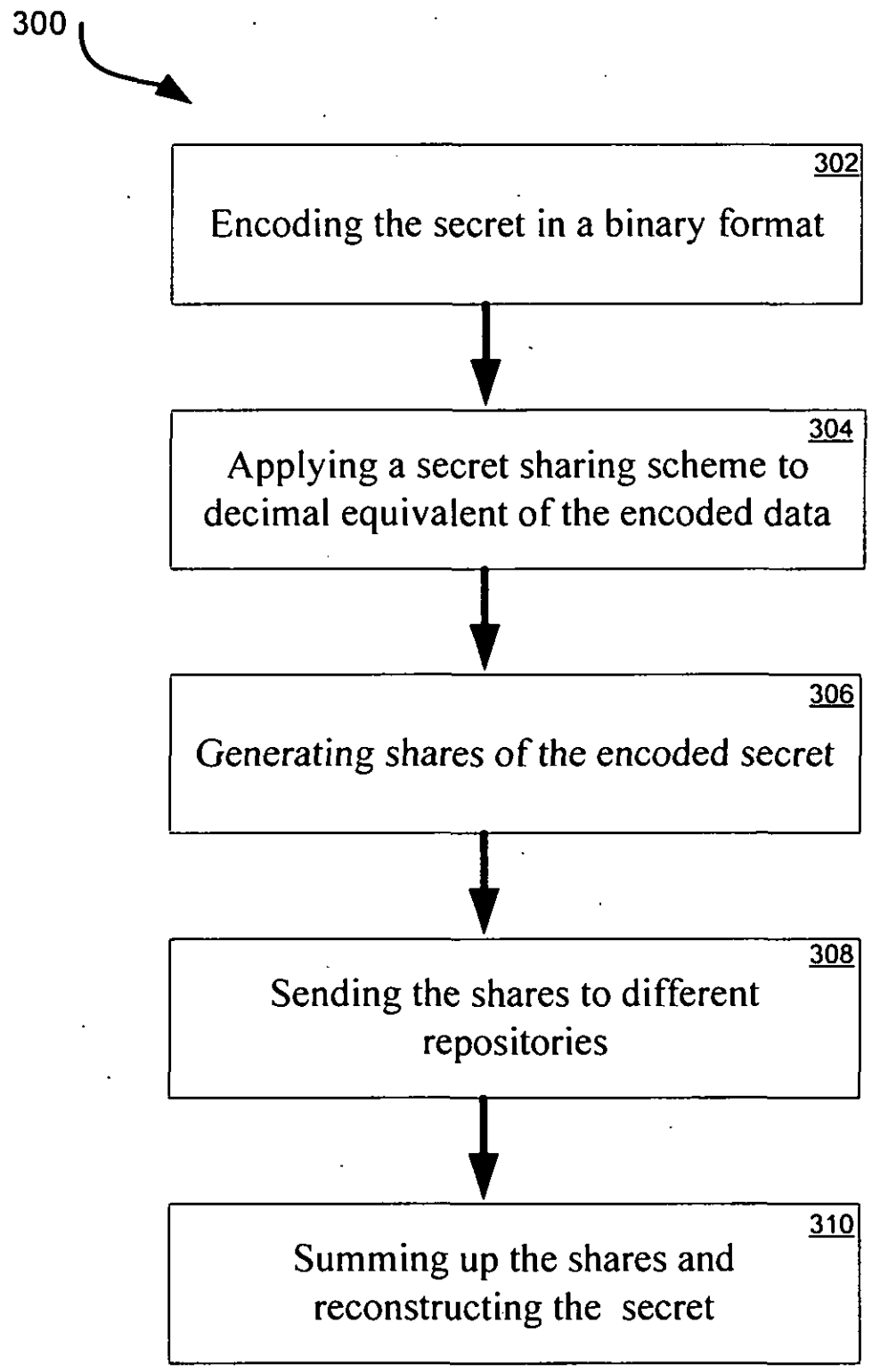


Figure 3

Mohammed Faisal (INPA No: 1941)
Head, IPR
L&T Technology Services Ltd.