(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 September 2014 (12.09.2014)

WIPO | PCT

(10) International Publication Number
WO 2014/137392 A1

(51) International Patent Classification:
H04L 9/00 (2006.01)

(21) International Application Number:
PCT/US2013/061690

(22) International Filing Date:
25 September 2013 (25.09.2013)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/772,404 4 March 2013 (04.03.2013) US

(71) Applicant *(for all designated States except US)*: **THOMSON LICENSING** [FR/FR]; 1-5, rue Jeanne d'Arc, F-92130 Issy Les Moulineaux (FR).

(72) Inventors; and
(71) Applicants : **NIKOLAENKO, Valeria** [US/US]; 63 Abrams Ct., Apt. 800, Stanford, California 94305 (US). **WEINSBERG, Udi** [IL/US]; 2135 Avy Ave, Menlo Park, California 94025 (US). **IOANNIDIS, Stratis** [GR/US]; 1961 Taylor Street, San Francisco, California 94133 (US). **JOYE, Marc** [BE/FR]; 22 rue du General Chazy, F-35300 Fougeres (FR). **TAFT, Nina** [US/US]; 4265 24th Street, San Francisco, California 94114 (US).

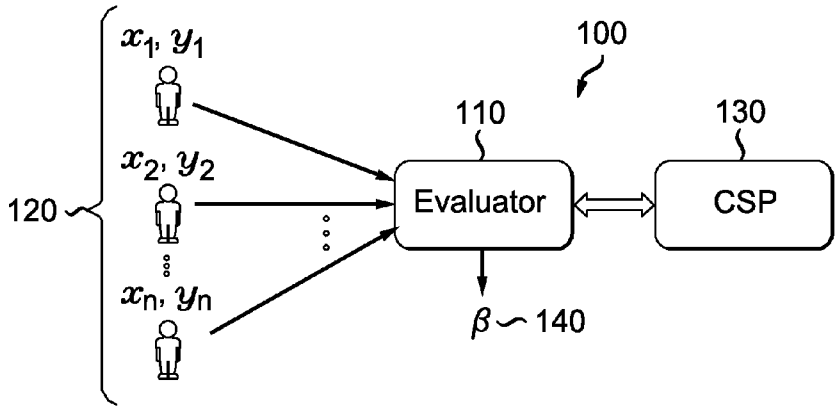
(74) Agents: **SHEDD, Robert D.** et al.; Thomson Licensing LLC, 2 Independence Way, Suite #200, Princeton, New Jersey 08540 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*

(54) **Title**: PRIVACY-PRESERVING RIDGE REGRESSION



(57) **Abstract**: A hybrid approach to privacy-preserving ridge regression is presented that uses both homomorphic encryption and Yao garbled circuits. Users in the system submit their data encrypted under a linearly homomorphic encryption. The linear homomorphism is used to carry out the first phase of the algorithm that requires only linear operations. The output of this phase generates encrypted data, in a form that is independent of the number of users n. In a second phase, a Yao garbled circuit that first implements homomorphic decryption and then does the rest of the regression algorithm (as shown, an optimized realization can avoid decryption in the garbled circuit) is evaluated. For this step a Yao garbled circuit approach is much faster than current fully homomorphic encryption schemes. Thus the best of both worlds is obtained by using linear homomorphisms to handle a large data set and using garbled circuits for the heavy non-linear part of the computation.

WO 2014/137392 A1

PRIVACY-PRESERVING RIDGE REGRESSION

REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of U.S. Provisional Application Serial No. 61/772,404 filed March 4, 2013 which is incorporated by reference herein in its entirety.

This application is also related to the applications entitled: "PRIVACY-PRESERVING RIDGE REGRESSION USING MASKS", and "PRIVACY-PRESERVING RIDGE REGRESSION USING PARTIALLY HOMOMORPHIC ENCRYPTION AND MASKS" which have been filed concurrently and are incorporated by reference herein in their entirety.

Background

Technical Field

The present invention generally relates to data mining and more specifically to protecting privacy during data mining using ridge regression.

Description of Related Art

Recommendation systems operate by collecting the preferences and ratings of many users for different items and running a learning algorithm on the data. The learning algorithm generates a model that can be used to predict how a new user will rate certain items. In particular, given the ratings that a user provides on certain items, the model can predict how that user will rate other items. There is a vast array of algorithms for generating such predictive models and many are actively used at large sites like Amazon and Netflix. Learning algorithms are also used on large medical databases, financial data, and many other domains.

In current implementations, the learning algorithm must see all user data in the clear in order to build the predictive model. In this disclosure it is determined whether the learning algorithm can operate without the data in the clear, thereby allowing users to retain control of their data. For medical data this allows for a model to be built without affecting user privacy. For books and movie preferences letting users keep control of their data reduces the risk of future unexpected embarrassment in case of a data breach at the service provider. Roughly

speaking, there are three existing approaches to data-mining private user data. The first lets users split their data among multiple servers using secret sharing. These servers then run the learning algorithm using a distributed protocol and privacy is assured as long as a majority of servers do not collude. The second is based on fully homomorphic encryption where the learning algorithm is executed over encrypted data and a trusted third party is trusted to only decrypt the final encrypted model. In a third approach Yao's garbled circuit construction could be used to compute on encrypted data and obtain a final model without learning anything else about user data. However an approach based upon Yao has never been applied to the regression class of algorithms before.

Summary

A hybrid approach to privacy-preserving ridge regression is presented that uses both homomorphic encryption and Yao garbled circuits. Users in the system submit their data encrypted under a linearly homomorphic encryption system such as Paillier or Regev. The Evaluator uses the linear homomorphism to carry out the first phase of the algorithm that requires only linear operations. This phase generates encrypted data. In this first phase, the system is asked to process a large number of records (proportional to the number of users in the system n). The processing in this first phase prepares the data such that the second phase of the algorithm is independent of n. In a second phase, the Evaluator evaluates a Yao garbled circuit that first implements homomorphic decryption and then does the rest of the regression algorithm (as shown, an optimized realization can avoid decryption in the garbled circuit). This step of the regression algorithm requires a fast linear system solver and is highly non-linear. For this step a Yao garbled circuit approach is much faster than current fully homomorphic encryption schemes. Thus the best of both worlds is obtained by using linear homomorphisms to handle a large data set and using garbled circuits for the heavy non-linear part of the computation. The second phase is also independent of n because of the way the computation is split into two phases.

In one embodiment method for privacy-preserving ridge regression is provided. The method includes the steps of requesting a garbled circuit from a crypto service provider; collecting data from multiple users that has been formatted and encrypted using homomorphic encryption; summing the data that has been formatted and encrypted using homomorphic encryption; and evaluating the garbled circuit from the crypto service provider with the summed data using oblivious transfer.

In another embodiment computing device for privacy-preserving ridge regression is provided. The computing device includes storage, memory, and a processor. The storage is for storing user data. The memory is for storing data for processing. The processor is configured to request a garbled circuit from a crypto service provider, collect data from multiple users that has been formatted and encrypted using homomorphic encryption, sum the data that has been formatted and encrypted using homomorphic encryption, and evaluate the garbled circuit from the crypto service provider with the summed data using oblivious transfer.

Objects and advantages will be realized and attained by means of the elements and couplings particularly pointed out in the claims. It is important to note that the embodiments disclosed are only examples of the many advantageous uses of the innovative teachings herein. It is to be understood that both the foregoing general description and the following detailed description are exemplary and explanatory and are not restrictive of the invention, as claimed. Moreover, some statements may apply to some inventive features but not to others. In general, unless otherwise indicated, singular elements may be in plural and vice versa with no loss of generality. In the drawings, like numerals refer to like parts through several views.

BRIEF SUMMARY OF THE DRAWINGS

FIGURE 1 depicts a block schematic diagram of a privacy-preserving ridge regression system according to an embodiment.

FIGURE 2 depicts a block schematic diagram of a computing device according to an embodiment.

FIGURE 3 depicts an exemplary garbled circuit according to an embodiment.

FIGURE 4 depicts a high level flow diagram of a methodology for providing a privacy-preserving ridge regression according to the embodiment.

FIGURE 5 depicts the operation of a first protocol for providing privacy-preserving ridge regression according to the embodiment.

FIGURE 6 depicts the operation of a first protocol for providing privacy-preserving ridge regression according to the embodiment.

FIGURE 7 depicts an exemplary embodiment of an algorithm for Cholesky decomposition according to the embodiment.

DETAILED DESCRIPTION

The focus of this disclosure is on a fundamental mechanism used in many learning algorithms, namely ridge regression. Given a large number of points in high dimension the regression algorithm produces a best-fit curve through these points. The goal is to perform the computation without exposing the user data or any other information about user data. This is achieved by using a system as shown in Figure 1:

In Figure 1, a block diagram of an embodiment of a system 100 for implementing privacy-preserving ridge regression is provided. The system includes an Evaluator 110, one or more users 120 and Crypto Service Provider (CSP) 130 which are in communication with each other. The Evaluator 110 is implemented on a computing device such as a server or personal computer (PC). The CSP 130 is similarly implemented on computing device such as a server or personal computer and is in communication with the Evaluator 110 over network, such as an Ethernet or Wi-Fi network. The one or more users 120 are in communication with the Evaluator 110 and CSP 130 via computing devices such as personal computers, tablets, smartphones, or the like.

Users 120 send encrypted data (from a PC, for example) to the Evaluator 110 (on a server, for example) which runs the learning algorithm. At certain points the Evaluator may interact with a Crypto Service Provider 130 (on another server) that is trusted not to collude with the Evaluator 110. The final outcome is the cleartext predictive model β 140.

Figure 2 depicts an exemplary computing device 200, such as a server, PC, tablet, or smartphone, that can be used to implement the various methodology and system elements for privacy-protecting ridge regression. The computing device 200 includes one or more processors 210, memory 220, storage 230, and a network interface 240. Each of these elements will be discussed in more detail below.

The processor 210 controls the operation of the electronic server 200. The processor 200 runs the software that operates the server as well as provides the functionality of cold start recommendations. The processor 210 is connected to memory 220, storage 230, and network interface 240, and handles the transfer and processing of information between these elements. The processor 210 can be general processor or a processor dedicated for a specific functionality. In certain embodiments there can be multiple processors.

The memory 220 is where the instructions and data to be executed by the processor are stored. The memory 210 can include volatile memory (RAM), non-volatile memory (EEPROM), or other suitable media.

The storage 230 is where the data used and produced the processor in executing the cold storage recommendation methodology of the present is stored. The storage may be magnetic media (hard drive), optical media (CD/DVD-Rom), or flash based storage.

The network interface 240 handles the communication of the server 200 with other devices over a network. An example of a suitable network is an Ethernet network. Other types of suitable home networks will be apparent to one skilled in the art given the benefit of this disclosure.

It should be understood that the elements set forth in Figure 2 are illustrative. The server 200 can include any number of elements and certain elements can provide part or all of the functionality of other elements. Other possible implementation will be apparent to on skilled in the art given the benefit of this disclosure.

SETTINGS AND THREAT MODEL

A. Architecture and Entities

Referring back to Figure 1, the system 100 is designed for many users 120 to contribute data to a central server called the Evaluator 110. The Evaluator 110 performs regression over the contributed data and produces a model, β 140, which can later be used for prediction or recommendation tasks. More specifically, each user $i = 1; : : : ; n$ has a private record comprising two variables $x_i \in \mathbb{R}^d$ and $y_i \in \mathbb{R}$, and the Evaluator wishes to compute $\beta \in \mathbb{R}^d$—the model —such that $y_i \backsimeq \beta^T x_i$. The goal is to ensure that the Evaluator learns nothing about the user's records beyond what is revealed by β 140, the final result of the regression algorithm. To initialize the system a third party is needed, which is referred the herein as a "Crypto Service Provider," that does most of its work offline.

More precisely, the parties in the system are the following, as shown in Figure 1.

- Users 120: each user i has private data x_i, y_i that it sends encrypted to the Evaluator 110.
- Evaluator 110: runs a regression algorithm on the encrypted data and obtains the learned model β 140 in the clear.

- Crypto Service Provider (CSP) 130: initializes the system 100 by giving setup parameters to the users 120 and the Evaluator 110.

The CSP 130 does most of its work offline long before the users 120 contribute their data to the Evaluator 110. In the most efficient design, the CSP 130 is also needed for a short one-round online step when the Evaluator 110 computes the model β 140.

B. Threat Model

The goal is to ensure that the Evaluator 110 and the CSP 130 cannot learn anything about the data contributed by users 120 beyond what is revealed by the final results of the learning algorithm. In the case that the Evaluator 110 colludes with some of the users 120, the users 120 should learn nothing about the data contributed by other users 120 beyond what is revealed by the results of the learning algorithm.

In this example, it is assumed that it is the Evaluator's 110 best interest to produce a correct model β 140. Hence, this embodiment is not concerned with a malicious Evaluator 110 which is trying to corrupt the computation in the hope of producing an incorrect result. However, the Evaluator 110 is motivated to misbehave and learn information about private data contributed by the users 120 since this data can potentially be sold to other parties, e.g., advertisers. Therefore, even a malicious Evaluator 110 should be unable to learn anything about user data beyond what is revealed by the results of the learning algorithm. The basic protocol which is only secure against an honest-but-curious Evaluator is set forth herein.

Non-threats: The system is not designed to defend against the following attacks:

- It is assumed that the Evaluator 110 and the CSP 130 do not collude. Each one may try to subvert the system as discussed above, but they do so independently. More precisely, when arguing security it is assumed that at most one of these two parties is malicious (this is an inherent requirement without which security cannot be achieved).
- It is assumed that the setup works correctly, that is all users 120 obtain the correct public key from the CSP 130. This can be enforced in practice with appropriate use of Certificate Authorities.

BACKGROUND

A. *Learning a Linear Model*

Briefly reviewing ridge regression, the algorithm that the evaluator 110 conducts in the system 110 to learn β 140. All results discussed below are classic, and can be found in most statistics and machine learning textbooks.

Linear Regression: Given a set of *n* input variables $x_i \in \mathbb{R}^d$, and a set of output variables $y_i \in \mathbb{R}$, the problem of learning a function $f: \mathbb{R}^d \to \mathbb{R}$ such that $y_i \backsimeq f(x_i)$ is known as *regression.* For example, the input variables could be a person's age, weight, body mass index, etc., while the output can be their likelihood to contract a disease.

Learning such a function from real data has many interesting applications that makes regression ubiquitous in data mining, statistics, and machine learning. On one hand, the function itself can be used for prediction, i.e., to predict the output value *y* of a new input $x \in \mathbb{R}^d$. Moreover, the structure of f can aid in identifying how different inputs affect the output—establishing, e.g., that weight, rather than age, is more strongly correlated to a disease.

Linear regression is based on the premise that *f* is well approximated by a linear map, i.e.,

$$y_i \backsimeq \beta^T x_i, \qquad i \in [n] \equiv \{1, \ldots, n\}$$

for some $\beta \in \mathbb{R}^d$. Linear regression is one of the most widely used methods for inference and statistical analysis in the sciences. In addition, it is a fundamental building block for several more advanced methods in statistical analysis and machine learning, such as kernel methods. For example, learning a function that is a polynomial of degree 2 reduces to linear regression over $x_{ik}x_{ik'}$, for $1 \leq k, k' \leq d$; the same principle can be generalized to learn any function spanned by a finite set of basis functions.

As mentioned above, beyond its obvious uses for prediction, the vector $\beta = (\beta_k)_{k=1,\ldots,d}$ is interesting as it reveals how *y* depends on the input variables. In particular, the sign of a coefficient β_k indicates either positive or negative correlation to the output, while the magnitude captures relative importance. To ensure these coefficients are comparable, but also for numerical stability, the inputs x_i are rescaled to the same, finite domain (e.g., [-1; 1]).

Computing the Coefficients: To compute the vector $\beta \in \mathbb{R}^d$, the latter is fit to the data by minimizing the following quadratic function over $\mathbb{R}^d$:

$$F(\beta) = \sum_{i=1}^{n} (y_i - \beta^T x_i)^2 + \lambda \|\beta\|_2^2 \ . \qquad (1)$$

The procedure of minimizing (1) is called *ridge regression*; the objective F(β) incorporates a penalty term $\lambda \| \beta \|_2^2$, which favors parsimonious solutions. Intuitively, for $\lambda = 0$, minimizing (1) corresponds to solving a simple least squares problem. For positive $\lambda > 0$, the term $\lambda \| \beta \|_2^2$ penalizes solutions with high norm: between two solutions that fit the data equally, one with fewer large coefficients is preferable. Recalling that the coefficients of β are indicators of how input affects output, this acts as a form of "Occam's razor": simpler solutions, with few large coefficients, are preferable. Indeed, a $\lambda > 0$ gives in practice better predictions over new inputs than the least squares solution based. Let $y \in \mathbb{R}^n$ be the vector of outputs and $x \in \mathbb{R}^{n \times d}$ be a matrix comprising the input vectors, one in each row; i.e.,

$$y = (y_i)_{i=1,\ldots,n} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

and

$$X = (x_i^T)_{i=1,\ldots,n} = \begin{pmatrix} x_{11} & x_{12} & \ldots & x_{1d} \\ x_{21} & x_{22} & \ldots & x_{2d} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \ldots & x_{nd} \end{pmatrix} \ .$$

The minimizer of (1) can be computed by solving the linear system

$$A\beta = b \qquad (2)$$

where $A = X^T X + \lambda I$ and $b = X^T y$. For $\lambda > 0$, the matrix A is symmetric positive definite, and an efficient solution can be found using the Cholesky decomposition as outlined below.

B. Yao's Garbled Circuits

In its basic version, Yao's protocol (a.k.a. garbled circuits) allows the two-party evaluation of a function $f(x_1; x_2)$ in the presence of semi-honest adversaries. The protocol is run between

the input owners (a_i denotes the private input of user i). At the end of the protocol, the value of $f(a_1; a_2)$ is obtained but no party learns more than what is revealed from this output value.

The protocol goes as follows. The first party, called *garbler*, builds a "garbled" version of a circuit computing f. The garbler then gives to the second party, called evaluator, the garbled circuit as well as the garbled-circuit input values that correspond to a_1 (and only those ones). The notation GI(a_1) is used to denote these input values. The garbler also provides the mapping between the garbled-circuit output values and the actual bit values. Upon receiving the circuit, the evaluator engages in a 1-out-of-2 oblivious transfer protocol with the garbler, playing the role of the chooser, so as to obliviously obtain the garbled-circuit input values corresponding to its private input a_2, GI(a_2). From GI(a_1) and GI(a_2), the evaluator can therefore calculate f($a_1; a_2$).

In more detail, the protocol evaluates the function f through a Boolean circuit 300 as seen in Figure 3. To each wire w_i *310,320* of the circuit, the garbler associates two random cryptographic keys, $K_{w_i}^0$ and $K_{w_i}^1$, that respectively correspond to the bit-values $b_i = 0$ and $b_i = 1$. Next, for each binary gate g (e.g., an OR-gate) with input wires (w_i,w_j)310, 320 and output wire w_k 330, the garbler computes the four ciphertexts

$$\mathrm{Enc}_{(K_{w_i}^{b_i}, K_{w_j}^{b_j})}(K_{w_k}^{g(b_i,b_j)}) \quad \text{for } b_i, b_j \in \{0,1\} \ .$$

The set of these four randomly ordered ciphertexts defines the garbled gate.

It is required that the symmetric encryption algorithm Enc, which is keyed by a pair of keys, has indistinguishable encryptions under chosen-plaintext attacks. It is also required that given the pair of keys $(K_{w_i}^{b_i}, K_{w_j}^{b_j})$, the corresponding decryption process unambiguously recovers the value of $K_{w_k}^{g(b_i,b_j)}$ from the four ciphertexts constituting the garbled gate. It is worth noting that the knowledge of $(K_{w_i}^{b_i}, K_{w_j}^{b_j})$ yields only the value of $K_{w_k}^{g(b_i,b_j)}$ and that no other output values can be recovered for this gate. So the evaluator can evaluate the entire garbled circuit gate-by-gate so that no additional information leaks about intermediate computations.

HYBRID APPROACH

Recall that, in this setup, each input and output variable $x_i,\ y_i, i \in [n]$, is private, and held by a different user. The Evaluator 110 wishes to learn the β determining the linear relationship between the input and output variables, as obtained through ridge regression with a given $\lambda > 0$.

As described in above, to obtain β, one needs the matrix $A \in \mathbb{R}^{d \times d}$ and the vector $b \in \mathbb{R}^{d}$, as defined in equation (2). Once these values are obtained, the Evaluator 110 can solve the linear system of equation (2) and extract β. There are several ways to tackle this problem in a privacy-preserving fashion. One can for example rely on secret sharing or on fully homomorphic encryption. Presently, these techniques seem to be unsuitable for the present setting as they lead to significant (on-line) communication or computation overhead. Consequently, Yao's approach is explored, as outlined in above.

One simple way to use Yao's approach is to design a single circuit with inputs x_i, y_i, for $i \in [n]$, and $\lambda > 0$, that computes the matrices A and b and subsequently solves the system $A\beta = b$. Such an approach has been used in the past for the computation of simple functions of inputs coming from multiple users, such the winner of an auction. Putting implementation issues aside (such as how to design a circuit that solves a linear system), a major shortcoming of such a solution is that the resulting garbled circuit depends on both the number of users n, as well as the dimension d of β and the input variables. In practical applications it is common that n is large, and can be in the order of millions of users. In contrast, d is relatively small, in the order of 10s. It is therefore preferable to reduce, or even eliminate, the dependency of the garbled circuit in n, so as to get a scalable solution. To this end, the problem was reformulated as discussed below.

A. Reformulating the Problem

Note that the matrix A and vector b can be computed in an iterative fashion, as follows. Assuming that each x_i and corresponding y_i are held by different users, each user i can locally compute the matrix $A_i = x_i x_i^T$ and the vector $b_i = y_i x_i$. It is then easily verified that summing the partial contributions yields:

$$A = \sum_{i=1}^{n} A_i + \lambda I \quad \text{and} \quad b = \sum_{i=1}^{n} b_i \ . \qquad (3)$$

Equation (3) importantly shows that A and b are the result of a series of additions. The Evaluator's regression task can therefore be separated into two subtasks: (a) collecting the A_i's and b_i's, to construct matrix A and vector b, and (b) using these to obtain β through the solution of the linear system (2).

Of course, the users cannot send their local shares, $(A_i; b_i)$, to the Evaluator in the clear. However, if the latter are encrypted using a public-key additive homomorphic encryption scheme, then the Evaluator 110 can reconstruct the encryptions of A and b from

the encryptions of the $(A_i; b_i)$'s. The remaining challenge is to solve equation (2), with the help of the CSP 130, without revealing (to the Evaluator 110 or the CSP 130) any additional information other than β; two distinct ways of doing so through the use of Yao's garbled circuits are described below.

More explicitly, let

$$\mathfrak{E}_{pk} : (A_i; b_i) \in \mathcal{M} \mapsto c_i = \mathfrak{E}_{pk}(A_i; b_i)$$

be a semantically secure encryption scheme indexed by a public key *pk* that takes on input a pair $(A_i; b_i)$ in the message space $\mathcal{M}$ and returns the encryption of $(A_i; b_i)$ under *pk*, c_i. Then it must hold for any pk and any two pairs $(A_i; b_i)$, $(A_j; b_j)$, that

$$\mathfrak{E}_{pk}(A_i; b_i) \otimes \mathfrak{E}_{pk}(A_j; b_j) = \mathfrak{E}_{pk}(A_i + A_j; b_i + b_j)$$

for some public binary operator. Such an encryption scheme can be constructed from any semantically secure additive homomorphic encryption scheme by encrypting component-wise the entries of A_i and b_i. Examples include Regev's scheme and Paillier's scheme.

Protocols are now ready to be presented. A high-level flow chart 400 is provided in Fig. 4. The flow chart 400 includes a preparation phase 410, a first phase (Phase 1) 420, and a second phase (Phase 2) 430. The phase of aggregating the user shares is referred to as *Phase 1* 420, and note that the addition it involves depends linearly in n. The subsequent phase, which amounts to computing the solution to Equation (2) from the encrypted values of A and b, is referred to as *Phase 2* 430. Note that Phase 2 430 has no dependence on n. These phases will be discussed below in conjunction with specific protocols. Note that it is assumed below the existence of a circuit that can solve the system $A\beta = b$; how such a circuit can be implemented efficiently is discussed in herein.

B. First Protocol

A high level depiction 500 of the operation of the first protocol can be seen in Figure 5. The first protocol operates as follows. As set forth above, the first protocol comprises three phases: a preparation phase 510, Phase 1 520, and Phase 2 530. As will become apparent, only Phase 2 530 really requires an on-line treatment.

Preparation phase(510). The Evaluator 110 provides the specifications to the CSP 130, such as the dimension of the input variables (i.e., parameter d) and their value range. The CSP 130 prepares a Yao garbled circuit for the circuit described in Phase 2 530 and

makes the garbled circuit available to the Evaluator 110. The CSP 130 also generates a public key pk_{csp} and a private key sk_{csp} for the homomorphic encryption scheme $\mathfrak{E}$, while the Evaluator 110 generates a public key pk_{ev} and a private key sk_{ev} for an encryption scheme $\mathcal{E}$ (that need not be homomorphic).

Phase 1 (520). Each user *i* locally computes her partial matrix A_i and vector b_i. These values are then encrypted using additive homomorphic encryption scheme $\mathfrak{E}$ under the public encryption key pk_{csp} of the CSP 130; i.e.,

$$c_i = \mathfrak{E}_{pk_{csp}}(A_i; b_i)$$

To prevent the CSP 130 from getting access to this value, the user *i* super-encrypts the value of c_i under the public encryption key pk_{ev} of the Evaluator 110 ; i.e.,

$$C_i = \mathcal{E}_{pk_{ev}}(c_i)$$

and sends C_i to the Evaluator 110.

The Evaluator 110 computes $c_\lambda = \mathfrak{E}_{pk_{csp}}(\lambda I; 0)$. It subsequently collects all received C_i's and decrypts them using its private decryption key sk_{ev} to recover the c_i's ; i.e.,

$$c_i = \mathcal{D}_{sk_{ev}}(C_i), \qquad \text{for } 1 \le i \le n$$

It then aggregates the so-obtained values and gets:

$$c = \left(\bigotimes_{i=1}^{n} c_i\right) \otimes c_\lambda = \mathfrak{E}_{pk_{csp}}\left(\sum_{i=1}^{n} A_i + \lambda I; \sum_{i=1}^{n} b_i\right)$$
$$= \mathfrak{E}_{pk_{csp}}(A; b) \quad . \qquad (4)$$

Phase 2 (530). The garbled circuit provided by the CSP 130 in the preparation phase 510 is a garbling of a circuit that takes as input GI(c) and does the following two steps:

1) decrypting c with sk_{csp} to recover *A* and *b* (here sk_{csp} is embedded in the garbled circuit); and

2) solving equation (2) and returning β.

In this Phase 2 530, the Evaluator 110 need only to obtain the garbled-circuit input values corresponding to c; i.e., GI(c). These are obtained using a standard Oblivious Transfer (OT) between the Evaluator 110 and the CSP 130.

The above hybrid computation performs a decryption of the encrypted inputs within the garbled circuit. As this can be demanding, it is suggested to use for example Regev homomorphic encryption scheme as the building block for $\mathfrak{E}$ since the Regev scheme has a very simple decryption circuit.

C. Second Protocol

A high level depiction 600 of the operation of the second protocol can be seen in Figure 6. The second protocol presents a modification that avoids decrypting $(A; b)$ in the garbled circuit using random masks. Phase 1 610 remains broadly the same. Thus Phase 2 will be highlighted (and the corresponding preparation phase). The idea is to exploit the homomorphic property to obscure the inputs with an additive mask. Note that if $(\mu_A; \mu_b)$ denotes an element in $\mathcal{M}$ (namely, the message space of homomorphic encryption $\mathfrak{E}$) then it follows from equation (4) that

$$c \otimes \mathfrak{E}_{\mathrm{pk}_{\mathrm{csp}}}(\mu_A; \mu_b) = \mathfrak{E}_{\mathrm{pk}_{\mathrm{csp}}}(A + \mu_A; b + \mu_b)$$

Hence assume that the Evaluator 110 chooses a random mask $(\mu_A; \mu_b)$ in $\mathcal{M}$, obscures c as above, and sends the resulting value to the CSP 130. Then, the CSP 130 can apply its decryption key and recover the masked values

$$\hat{A} = A + \mu_A \text{ and } \hat{b} = b + \mu_b$$

As a consequence, one can apply the protocol of the previous section where the decryption is replaced by the removal of the mask. In more detail, it involves:

Preparation phase (610). As before, the Evaluator 110 sets up the evaluation. The Evaluator 110 provides the specifications to the CSP 130 to build a garbled circuit supporting its evaluation. The CSP 130 prepares the circuit and makes it available to the Evaluator 110, and both generate public and private keys. The Evaluator 110 chooses a random mask $(\mu_A; \mu_b) \in \mathcal{M}$ and engages in an Oblivious Transfer (OT) protocol with the CSP 130 to get the garbled-circuit input values corresponding to $(\mu_A; \mu_b)$; i.e., GI $\mu_A; \mu_b$).

Phase 1 (620). This is similar to the first protocol. In addition, the Evaluator 110masks c as

$$\hat{c} = c \otimes \mathfrak{E}_{pk_{csp}}(\mu_A; \mu_b)$$

Phase 2 (630). The Evaluator 110 sends $\hat{c}$ to the CSP 130 that decrypts it to obtain $(\hat{A}; \hat{b})$ in the clear. The CSP 130 then sends the garbled input values GI($\hat{A}$; $\hat{b}$) back to the Evaluator 110. The garbled circuit provided by the CSP 130 in the preparation phase is a garbling of a circuit that takes as input GI($\hat{A}$; $\hat{b}$) and GI(μ_A; μ_b) and does the following two steps:

1) subtracts the mask (μ_A; μ_b) from ($\hat{A}$; $\hat{b}$) to recover A and b;

2) solves equation (2) and returns β.

The garbled circuit as well as the garbled-circuit input values corresponding to (μ_A; μ_b), GI(μ_A; μ_b), were obtained during the preparation phase 610. In this phase, the Evaluator 110 need only receive from the CSP 130 the garbled circuit input values corresponding to ($\hat{A}$; $\hat{b}$), GI($\hat{A}$; $\hat{b}$). Note that there is no Oblivious Transfer (OT) in this phase.

For this second realization, the decryption is not executed as part of the circuit. Therefore one is not restricted to selecting a homomorphic encryption scheme that can be efficiently implemented as a circuit. Instead of Regev's scheme, it is suggested to use Paillier's scheme or its generalization by Damgård and Jurik as the building block for $\mathfrak{E}$. These schemes have a shorter ciphertext expansion than Regev and require smaller keys.

D. Third Protocol

For some applications, a related idea applies when the homomorphic encryption scheme has only a partial homomorphic property. This notion is made explicit in the next definition.

Definition 1: A *partially homomorphic encryption scheme* is an encryption scheme such that it is possible to add (if the partial homomorphism is additive) or to multiply (if the partial homomorphism is multiplicative) a constant to an encrypted plaintext without needing the private encryption key.

Here are some examples.

- Let $\mathbb{F}_p$denote a prime field and let G = $\langle g \rangle$be a cyclic subgroup of the multiplicative group $\mathbb{F}_p^*$, generated by g. Let q denote the order of G. For plain ElGamal encryption, the message space is $\mathcal{M} = G$. The public encryption key is $y = g^x$ while the private key

is x. The encryption of a message m in $\mathcal{M}$ is given by $(R; c)$ with $R = g^r$ and $c = my^r$ for some random $r \in \mathbb{Z}/q\mathbb{Z}$. Plaintext m is then recovered using secret key x as $m = c/R^x$.

– The above system is partially homomorphic with respect to the multiplication in $\mathbb{F}_p^*$: For any constant $K \in \mathcal{M}$, $C' = (R; Kc)$ is the encryption of message $m' = Km$.

- The so-called hashed ElGamal cryptosystem requires in addition an hash function H, mapping group elements from G to $\mathbb{F}_2^k$, , for some parameter k. The message space is $\mathcal{M} = \mathbb{F}_2^k$. The key generation is as for plain ElGamal. The encryption of a message $m \in \mathcal{M}$ is given by $(R; c)$ with $R = g^r$ and $c = m + H(y^r)$ for some random $r \in \mathbb{Z}/q\mathbb{Z}$. Plaintext m is then recovered using secret key x as $m = c+H(R^x)$. Note that '+' corresponds to the addition in $\mathbb{F}_2^k$ (i.e., it can equivalently be seen as an XOR on k-bit strings).

  – The above system is partially homomorphic with respect to the XOR: For any constant $K \in \mathcal{M}$, $C' = (R;K + c)$ is the encryption of message $m' = K + m$.

For the sake of non-limiting example, suppose now that c is the encryption of $(A; b)$ under a partially homomorphic encryption scheme, say $\mathfrak{E}$, then if $(\mu_A; \mu_b)$ denotes an element in $\mathcal{M}$ (namely, the message space of partially homomorphic encryption $\mathfrak{E}$) then it follows from equation (4) that

$$c \oplus \mathfrak{E}_{\mathrm{pk}_{\mathrm{csp}}}(\mu_A; \mu_b) = \mathfrak{E}_{\mathrm{pk}_{\mathrm{csp}}}(A + \mu_A; b + \mu_b)$$

for some operator $\oplus$. (In the above description, the homomorphism is noted additively; the same holds true for a multiplicatively written homomorphism.)

Hence, assume that the Evaluator 110 chooses a random mask $(\mu_A; \mu_b)$ in $\mathcal{M}$, obscures c as above, and sends the resulting value to the CSP 130. Then, the CSP 130 can apply its decryption key and recover the masked values

$$\hat{A} = A + \mu_A \text{ and } \hat{b} = b + \mu_b$$

As a consequence, the protocol of the previous section can be applied where the decryption is replaced by the removal of the mask.

Finally, note that the trick of using a mask as per the second or third protocol is not limited to the case of ridge regression. It can be used in any application combining in a hybrid

way homomorphic encryption (respectively partially homomorphic encryption) with garbled circuits.

E. Discussion

The proposed protocols have several strengths that make them efficient and practical in real-world scenarios. First, there is no need for users to stay on-line during the process. Since Phase 1 420 is incremental, each user can submit their encrypted inputs, and leave the system.

Furthermore, the system 100 can be easily applied to performing ridge regression multiple times. Assuming that the Evaluator 110 wishes to perform ℓ estimations, it can retrieve ℓ garbled circuits from the CSP 130 during the preparation phase 410. Multiple estimations can be used to accommodate the arrival of new users 120. In particular, since the public keys are long-lived, they do not need to be refreshed too often, meaning that when new users submit more pairs $(A_i; b_i)$ to the Evaluator 110, the latter can sum them with the prior values and compute an updated β. Although this process requires utilizing a new garbled circuit, the users that have already submitted their inputs do not need to resubmit them.

Finally, the amount of required communications is significantly smaller than in a secret sharing scheme, and only the Evaluator 110 and the CSP 130 communicate using Oblivious Transfer (OT). Note also that, rather than using the public key encryption scheme $\mathcal{E}$ in Phase 1 420, the users can use any means to establish a secure communication with the Evaluator 110, such as, e.g., SSL.

F. Further Optimizations

Recall that the matrix A is in $\mathbb{R}^{d\times d}$ and the vector b is in $\mathbb{R}^d$. Hence letting k denote the bit-size used to encode real numbers, the matrix A and vector b respectively need d^2k bits and dk bits for their representation. The second protocol requires a random mask $(\mu_A; \mu_b)$ in $\mathcal{M}$. Suppose that the homomorphic encryption scheme $\mathfrak{E}$ was built on top of Paillier's scheme where every entry of A and of b is individually Paillier encrypted. In this case the message space $\mathcal{M}$ of $\mathfrak{E}$ is composed of $(d^2 + d)$ elements in $\mathbb{Z}/N\mathbb{Z}$ for some RSA modulus N. But as those elements are k-bit values there is no need to draw the corresponding masking values in the whole range $\mathbb{Z}/N\mathbb{Z}$. Any $(k+l)$-bit values for some (relatively short) security length l will do, as long as they statistically hide the corresponding entry. In practice, this leads to fewer Oblivious Transfers in the preparation phase and to a smaller garbled circuit.

Another way to improve the efficiency is via a standard batching technique, that is packing multiple plaintext entries of A and b into a single Paillier ciphertext. For example, packing 20 plaintext values into a single Paillier ciphertext (separated by sufficiently many 0's) will reduce the running time of Phase 1 by a factor of 20.

IMPLEMENTATION

To assess the practicality of the privacy-preserving system, the system was implemented and tested on both synthetic and real datasets. The second protocol proposed above was implemented, as it does not require decryption within the garbled circuit, and allows for the use of homomorphic encryption that is efficient for Phase 1 (that only involves summation).

A. Phase 1 Implementation

As discussed above, for homomorphic encryption Paillier's scheme was use with a 1024 bits long modulus, which corresponds to 80-bits security level. To speed up Phase 1, batching was also implemented as outlined in above. Given n users that contribute their inputs, the number of elements that can be batched into one Paillier ciphertext of 1024 bits is $1024=(b + \log_2 n)$, where b is the total number of bits for representing numbers. As discussed later, b is determined as a function of the desired accuracy, thus in this experiment, between 15 and 30 elements were batched.

B. Circuit Garbling Framework

The system was built on top of FastGC, a Java-based open-source framework that enables developers to define arbitrary circuits using elementary XOR, OR and AND gates. Once the circuits are constructed, the framework handles garbling, oblivious transfer and the complete evaluation of the garbled circuit. FastGC includes several optimizations. First, the communication and computation cost for XOR gates in the circuit is significantly reduced using the "free XOR" technique. Second, using the garbled-row reduction technique, FastGC reduces the communication cost for k-fan-in non-XOR gates by $1=2^k$, which gives a 25% communication saving, since only 2-fan-in gates are defined in the framework. Third, FastGC implements the OT extension which can execute a practically unlimited number of transfers at the cost of k OTs and several symmetric-key operations per additional OT. Finally, the last optimization is the succinct "addition of 3 bits" circuit, which defines a circuit with four XOR gates (all of which are "free" in terms of communication and computation) and just one AND gate. FastGC enables the garbling and evaluation to take place concurrently. More

specifically, the CSP 130 transmits the garbled tables to the Evaluator 110 as they are produced in the order defined by circuit structure. The Evaluator 110 then determines which gate to evaluate next based on the available output values and tables. Once a gate was evaluated its corresponding table is immediately discarded. This amounts to the same computation and communication costs as pre-computing all garbled circuits off-line, but brings memory consumption to a constant.

C. Solving a Linear System in a Circuit

One of the main challenges of the present approach is designing a circuit that solves the linear system $A\beta = b$, as defined in equation (2). When implementing a function as a garbled circuit, it is preferable to use operations that are data-agnostic, i.e., whose execution path does not depend on the input. For example, as inputs are garbled, the Evaluator 110 needs to execute all possible paths of an if-then-else statement, which leads to an exponential growth of both the circuit size and the execution time in the presence of nested conditional statements. This renders impractical any of the traditional algorithms for solving linear systems that require pivoting, such as, e.g., Gaussian elimination.

For the sake of simplicity, this system implemented the standard Cholesky algorithm presented below. Note, however, that its complexity can be further reduced to the same complexity as block-wise inversion using similar techniques.

There are several possible decomposition methods for solving linear systems. Cholesky decomposition is a data-agnostic method for solving a linear system that is applicable only when the matrix A is symmetric positive definite. The main advantage of Cholesky is that it is numerically robust without the need for pivoting. In particular, it is well suited for fixed point number representations.

Since $A = \lambda I + \sum_i^n x_i x_i^T$ is indeed a positive definite matrix for $\lambda > 0$, Cholesky was chosen as the method of solving $A\beta = b$ in this implementation.

The main steps of Cholesky decomposition are briefly outlined below. The algorithm constructs a lower-triangular matrix L such that $A = L^T L$. Solving the system $A\beta = b$ then reduces to solving the following two systems:

$$L^T y = b \text{ ; and}$$

$$L\beta = y$$

Since matrices L and LT are triangular, these systems can be solved easily using back substitution. Moreover, because matrix A is positive definite, matrix L necessarily has nonzero values on the diagonal, so no pivoting is necessary.

The decomposition $A = L^T L$ is described in Algorithm 1 shown in Figure 7. It involves $\Theta(d^3)$ additions, $\Theta(d^3)$ multiplications, $\Theta(d^2)$divisions and $\Theta(d)$ square root operations. Moreover, the solution of the two systems above through backwards elimination involves $\Theta(d^2)$additions, $\Theta(d^2)$multiplications and $\Theta(d)$divisions. The implementation of these operations as circuits are discussed below.

D. Representing Real Numbers

In order to solve the linear system (2), it is necessary to accurately represent real numbers in a binary form. Two possible approaches for representing real Numbers were considered: floating point and fixed point. Floating point representation of a real number a is given by formula:

$$[a] = [m; p] \text{ ; where } a \approx 1.m \cdot 2^p$$

Floating point representation has the advantage of accommodating numbers of practically arbitrary magnitude. However, elementary operations on floating point representations, such as addition, are difficult to implement in a data-agnostic way. Most importantly, using Cholesky warrants using fixed point representation, which is significantly simpler to implement. Given a real number a, its fixed point representation is given by:

$$[a] = \lfloor a \cdot 2^p \rfloor \text{ , where the exponent } p \text{ is fixed.}$$

As discussed herein, many of the operations needed to be performed can be implemented in a data-agnostic fashion over fixed point numbers. As such, the circuits generated for fixed point representation are much smaller. Moreover, recall that the input variables of ridge regression xi are typically rescaled the same domain (between -1 and 1) to ensure that the coefficients of β are comparable, and for numerical stability. In such a setup, it is known that Cholesky decomposition can be performed on A with fixed point numbers without leading to overflows. Moreover, given bounds on y_i and the condition number of the matrix A, the bits necessary to prevent overflows can be computed while solving the last two triangular systems in the method. Thus the system was implemented using fixed point representations. The number of bits p for the fractional part can be selected as a system parameter, and creates a trade-off between the accuracy of the system and size of the generated circuits. However,

selecting p can be done in a principled way based on the desired accuracy. Negative numbers are represented using the standard two's complement representation.

The various embodiments disclosed herein can be implemented as hardware, firmware, software, or any combination thereof. Moreover, the software is preferably implemented as an application program tangibly embodied on a program storage unit or computer readable medium. The application program may be uploaded to, and executed by, a machine comprising any suitable architecture. Preferably, the machine is implemented on a computer platform having hardware such as one or more central processing units ("CPUs"), a memory, and input/output interfaces. The computer platform may also include an operating system and microinstruction code. The various processes and functions described herein may be either part of the microinstruction code or part of the application program, or any combination thereof, which may be executed by a CPU, whether or not such computer or processor is explicitly shown. In addition, various other peripheral units may be connected to the computer platform such as an additional data storage unit and a printing unit.

All examples and conditional language recited herein are intended for pedagogical purposes to aid the reader in understanding the principles of the embodiments and the concepts contributed by the inventor to furthering the art, and are to be construed as being without limitation to such specifically recited examples and conditions. Moreover, all statements herein reciting principles, aspects, and varies embodiments of the invention, as well as specific examples thereof, are intended to encompass both structural and functional equivalents thereof. Additionally, it is intended that such equivalents include both currently known equivalents as well as equivalents developed in the future, i.e., any elements developed that perform the same function, regardless of structure.

Claims:

1. A method for providing privacy-preserving ridge regression, the method comprising:

requesting a garbled circuit from a crypto service provider;

collecting data from multiple users that has been formatted and encrypted using homomorphic encryption;

summing the data that has been formatted and encrypted using homomorphic encryption; and

evaluating the garbled circuit from the crypto service provider with the summed data using oblivious transfer.

2. The method of claim 1, wherein the step of requesting a garbled circuit from a crypto service provider comprises:

providing a dimension of the input variables for the garbled circuit; and

providing the value range of the input variables.

3. The method of claim 1 wherein an evaluator implemented on a computing device performs the method.

4. The method of claim 3 wherein the crypto service provider is implemented on a computing device remote from the computing device the evaluator is implemented on.

5. The method of claim 1 further comprising the step of providing an encryption key for encrypting the data from multiple users.

6. The method of claim 5 wherein the data from multiple users is further encrypted with an encryption key provided by the crypto service provider.

7. The method of claim 1 wherein the step of evaluating the garbled circuit further comprises:

decrypting the summed data; and

solving the ridge regression equation embodied by the garbled circuit.

8. The method of claim 1 wherein the step of collecting data from multiple users comprises receiving data sent from each of the multiple users via a computing device.

9. A computing device for providing privacy-preserving ridge regression, the computer device comprising:

a storage for storing user data;

a memory for storing data for processing; and

a processor configured to request a garbled circuit from a crypto service provider, collect data from multiple users that has been formatted and encrypted using homomorphic encryption, sum the data that has been formatted and encrypted using homomorphic encryption, and evaluate the garbled circuit from the crypto service provider with the summed data using oblivious transfer.

10. The computing device of claim 9 further comprising a network connection for connecting to a network.

11. The computing device of claim 9 wherein the crypto service provider is implemented on a separate computing device.

12. The computing device of claim 9 wherein the step of requesting a garbled circuit from a crypto service provider comprises:

providing a dimension of the input variables for the garbled circuit; and

providing the value range of the input variables.

13. The computing device of claim 9 wherein the step of evaluating the garbled circuit further comprises:

decrypting the summed data; and

solving the ridge regression equation embodied by the garbled circuit.

14. The computing device of claim 9, wherein the data from multiple users is encrypted with an encryption key provided by the crypto service provider and encrypted with and encryption key by the computing device.

15. A machine readable medium containing instructions that when executed perform the steps comprising:

requesting a garbled circuit from a crypto service provider;

collecting data from multiple users that has been formatted and encrypted using homomorphic encryption;

summing the data that has been formatted and encrypted using homomorphic encryption; and

evaluating the garbled circuit from the crypto service provider with the summed data using oblivious transfer.

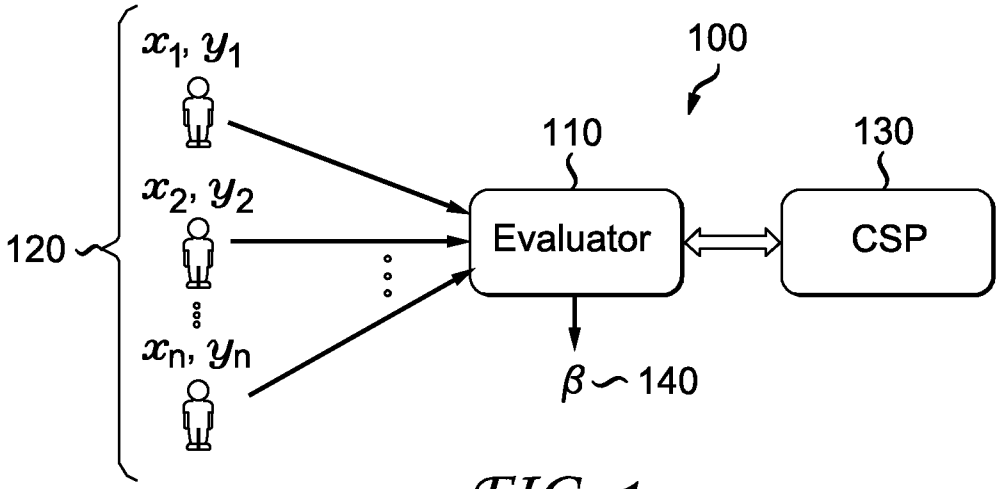

x_1, y_1
x_2, y_2
x_n, y_n
120
100
110
130
Evaluator
CSP
β
140

FIG. 1

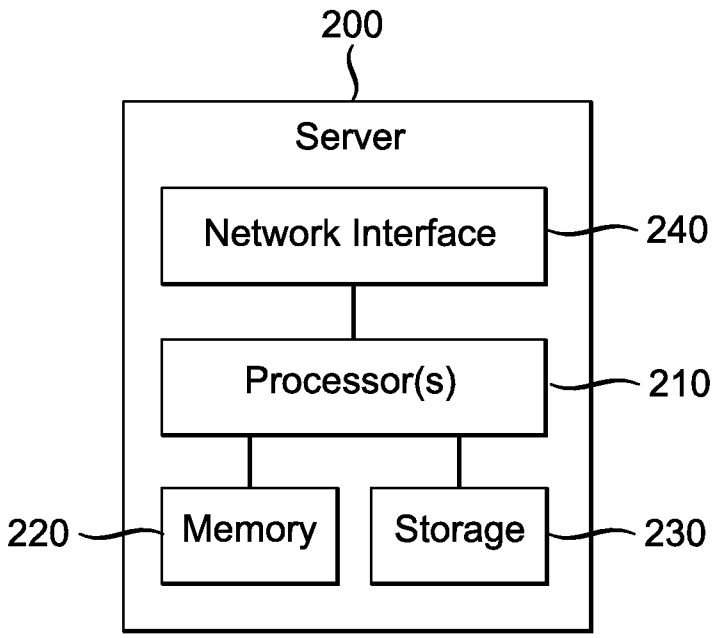

200
Server
Network Interface
240
Processor(s)
210
220
Memory
Storage
230

FIG. 2

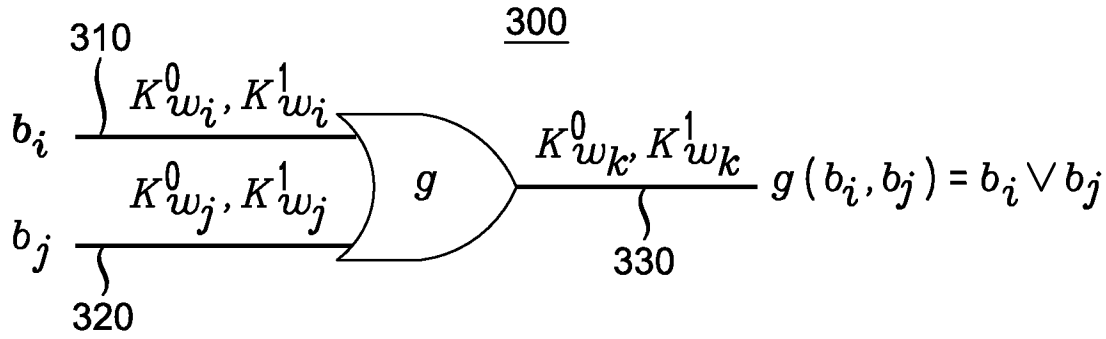

300
310
$K^0_{w_i}, K^1_{w_i}$
b_i
$K^0_{w_j}, K^1_{w_j}$
b_j
320
g
$K^0_{w_k}, K^1_{w_k}$
330
$g(b_i, b_j) = b_i \vee b_j$

FIG. 3

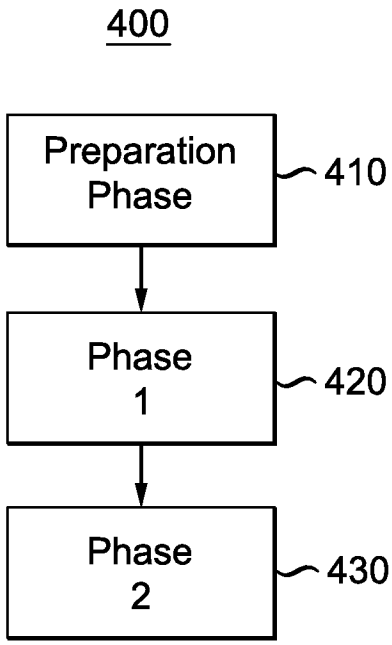

400
Preparation Phase
410
Phase 1
420
Phase 2
430

FIG. 4

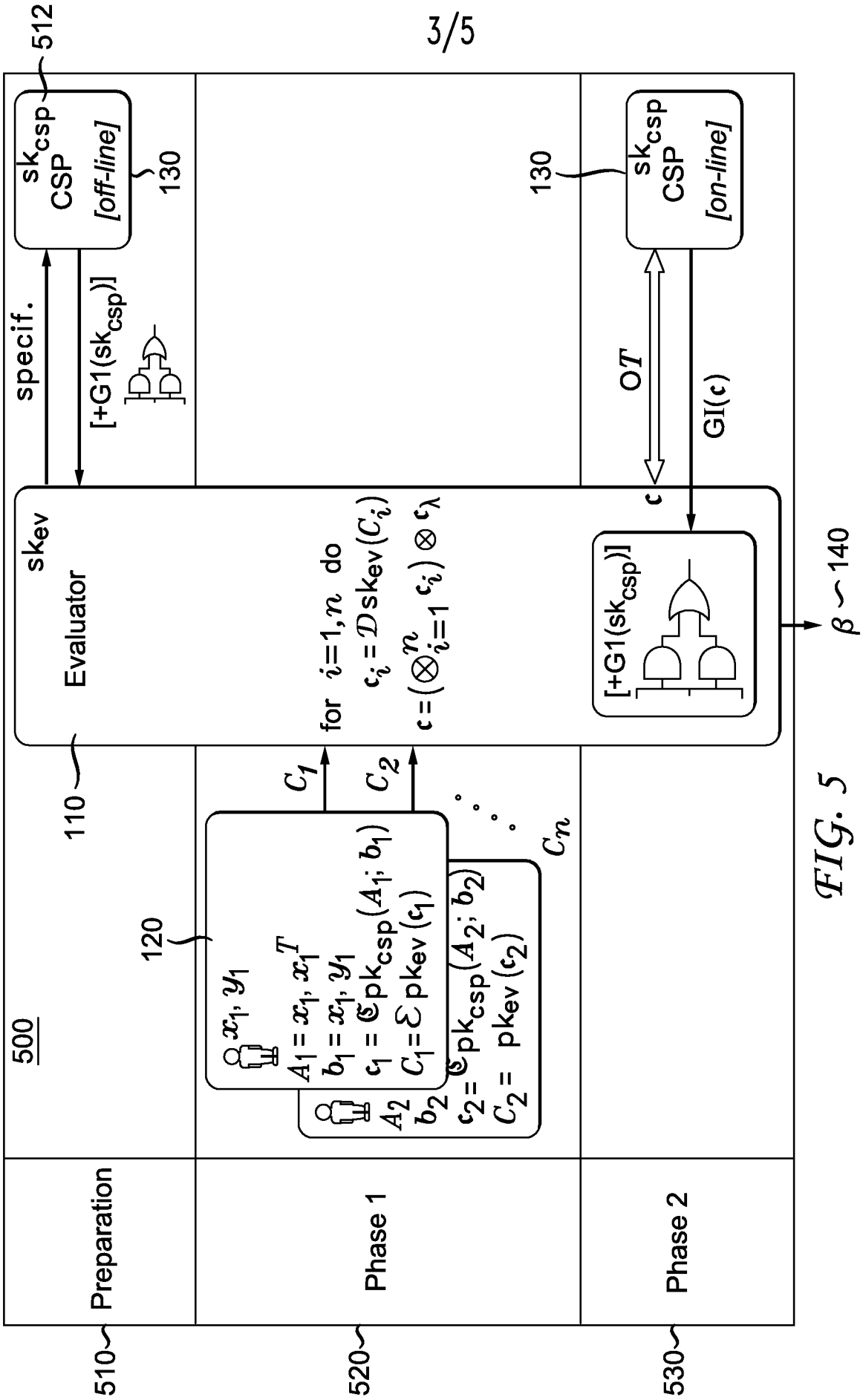

512
sk_{csp}
CSP
[off-line]
130
specif.
[+G1(sk_{csp})]
130
sk_{csp}
CSP
[on-line]
OT
GI(c)
sk_{ev}
Evaluator
110
for $i=1,n$ do
$c_i = \mathcal{D}sk_{ev}(C_i)$
$c = (\otimes_{i=1}^{n} c_i) \otimes c_\lambda$
c
[+G1(sk_{csp})]
β 140
C_1
C_2
C_n
120
x_1, y_1
$A_1 = x_1, x_1^T$
$b_1 = x_1, y_1$
$c_1 = \mathcal{E}pk_{csp}(A_1; b_1)$
$C_1 = \mathcal{E}pk_{ev}(c_1)$
A_2
b_2
$c_2 = \mathcal{E}pk_{csp}(A_2; b_2)$
$C_2 = \mathcal{E}pk_{ev}(c_2)$
500
Preparation
Phase 1
Phase 2
510
520
530

FIG. 5

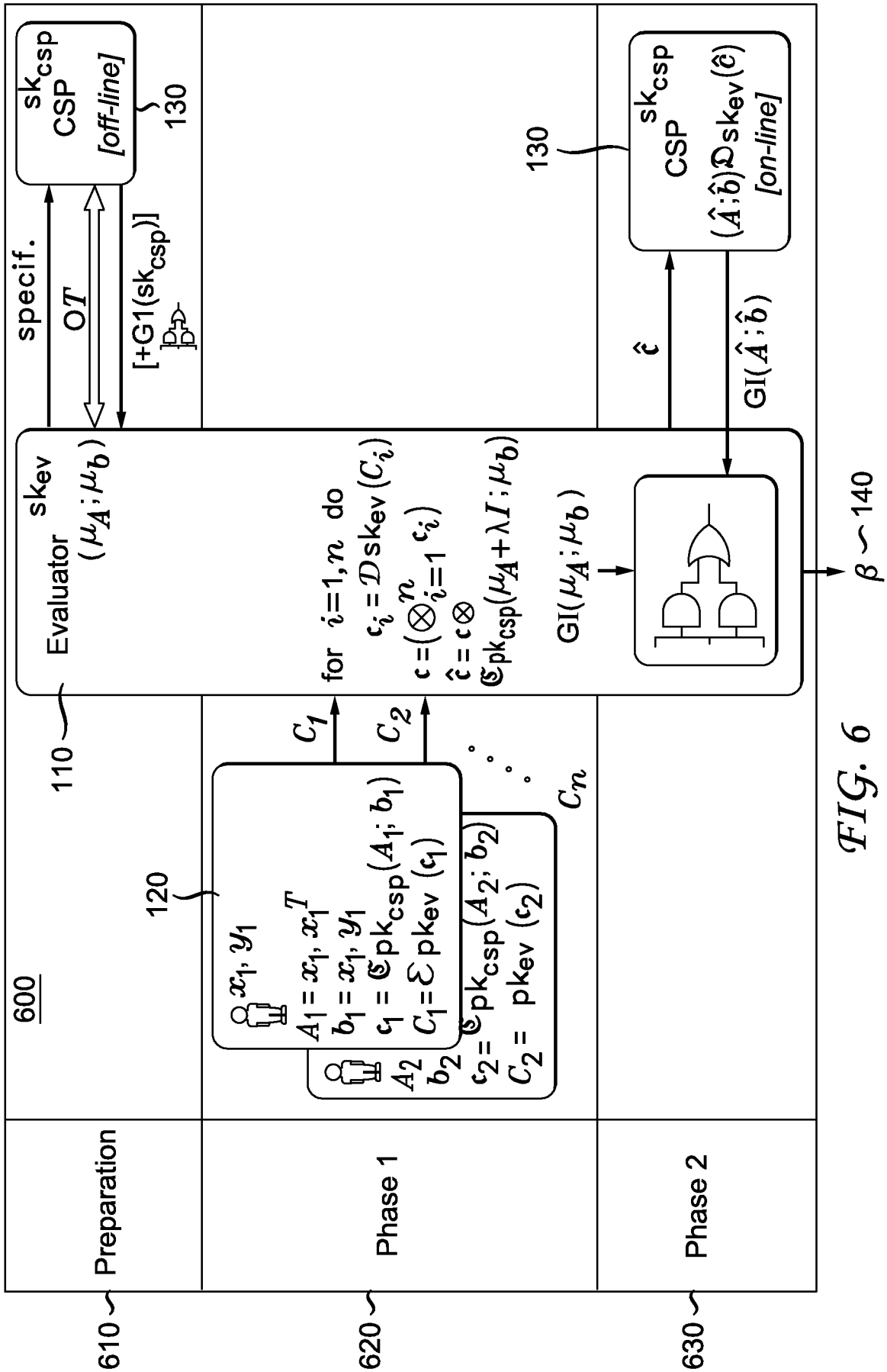

600
Preparation
Phase 1
Phase 2
610
620
630
CSP
[off-line]
130
specif.
OT
[+G1(sk_csp)]
Evaluator
sk_ev
(μ_A; μ_b)
110
for i=1,n do
c_i = D sk_ev (C_i)
c = (⊗_{i=1}^{n} c_i)
ĉ = c ⊗ E pk_csp(μ_A + λI; μ_b)
GI(μ_A; μ_b)
130
CSP
sk_csp
(Â; b̂) D sk_ev(ĉ)
[on-line]
ĉ
GI(Â; b̂)
β
140
120
x_1, y_1
A_1 = x_1, x_1^T
b_1 = x_1, y_1
c_1 = E pk_csp(A_1; b_1)
C_1 = E pk_ev (c_1)
A_2
b_2
c_2 = E pk_csp(A_2; b_2)
C_2 = pk_ev (c_2)
C_1
C_2
C_n

FIG. 6

Algorithm 1 The Cholesky decomposition of A $\in \mathbb{R}^{d \times d}$

Input: $A = (a_{ij})_{i,j \in \{1,\ldots,d\}}$

Output: L, s.t., $A = L^T L$

1: **for** j = 1, d **do**

2: **for** k = 1, $j - 1$ **do**

3: **for** $i = j$, d **do**

4: $a_{ij} -= a_{ik} a_{jk}$

5: $a_{jj} = \sqrt{a_{jj}}$

6: **for** $k = j + 1$, d **do**

7: $akj \ /= \ ajj$

8: **return** $L := A$

FIG. 7

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2013/061690

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04L9/00
ADD.

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
H04L

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

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	VALERIA NIKOLAENKO ET AL: "Privacy-Preserving Ridge Regression on Hundreds of Millions of Records", SECURITY AND PRIVACY (SP), 2013 IEEE SYMPOSIUM ON, IEEE, 19 May 2013 (2013-05-19), pages 334-348, XP032431333, DOI: 10.1109/SP.2013.30 ISBN: 978-1-4673-6166-8 the whole document ----- -/--	1-15

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
3 December 2013	11/12/2013

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Baltersee, Jens

2

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	AHMAD-REZA SADEGHI ET AL: "Efficient Privacy-Preserving Face Recognition", 2 December 2009 (2009-12-02), INFORMATION, SECURITY AND CRYPTOLOGY Â ICISC 2009, SPRINGER BERLIN HEIDELBERG, BERLIN, HEIDELBERG, PAGE(S) 229 - 244, XP019147094, ISBN: 978-3-642-14422-6 page 231, line 1 - line 8 Section 3.2 - Projection; page 235 page 237 figure 3(a) -----	1-15
A	TOMMASO PIGNATA ET AL: "General function evaluation in a STPC setting via piecewise linear approximation", INFORMATION FORENSICS AND SECURITY (WIFS), 2012 IEEE INTERNATIONAL WORKSHOP ON, IEEE, 2 December 2012 (2012-12-02), pages 55-60, XP032309458, DOI: 10.1109/WIFS.2012.6412625 ISBN: 978-1-4673-2285-0 Section II.C, "Hybrid protocols"; page 56 -----	1-15