Innovation, Sciences et Développement économique Canada
Office de la Propriété Intellectuelle du Canada

Innovation, Science and Economic Development Canada
Canadian Intellectual Property Office

CA 3108956 C 2023/09/05

(11)(21) 3 108 956

(12) BREVET CANADIEN
CANADIAN PATENT

(13) C

(22) **Date de dépôt/Filing Date:** 2021/02/11

(41) **Mise à la disp. pub./Open to Public Insp.:** 2021/04/22

(45) **Date de délivrance/Issue Date:** 2023/09/05

(30) **Priorité/Priority:** 2020/02/11 (US62/975,160)

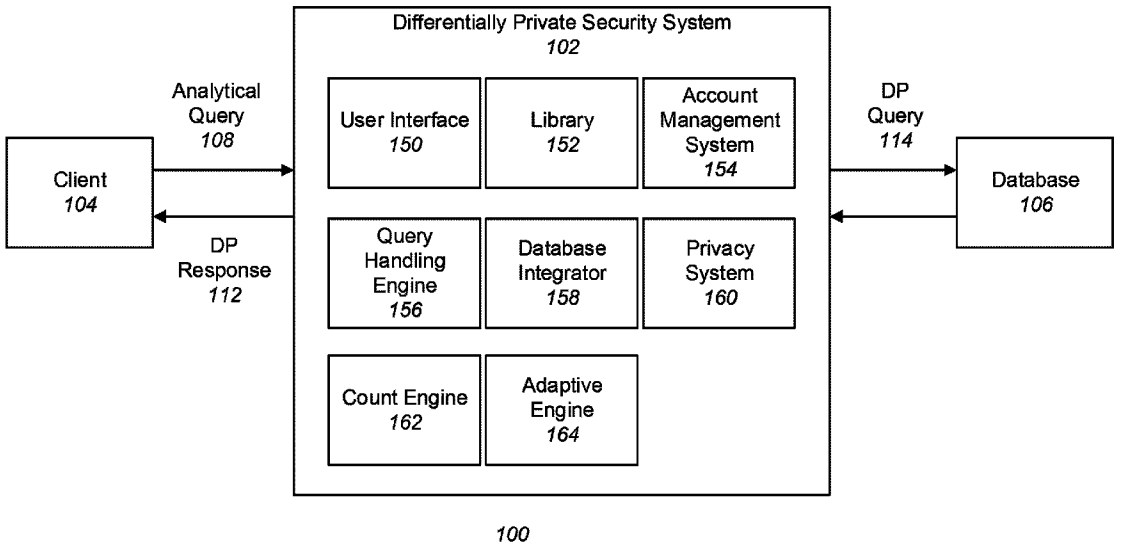
(51) **Cl.Int./Int.Cl.** ***G06F 16/90*** (2019.01),
G06F 16/903 (2019.01), ***G06F 17/18*** (2006.01),
G06F 21/60 (2013.01)

(72) **Inventeurs/Inventors:**
YANG, ANN, US;
DAMEWOOD, LIAM, US;
NICULAESCU, OANA, US;
ROZENSHTEYN, ALEXANDER, US

(73) **Propriétaire/Owner:**
SNOWFLAKE INC., US

(74) **Agent:** GOWLING WLG (CANADA) LLP

(54) **Titre : COMPTE DIFFERENTIELLEMENT PRIVE ADAPTATIF**
(54) **Title: ADAPTIVE DIFFERENTIALLY PRIVATE COUNT**



(57) **Abrégé/Abstract:**
A differentially private security system communicatively coupled to a database storing restricted data receives a database query from a client. The database query includes an operation, a target accuracy, and a maximum privacy spend for the query. The system performs the operation to produce a result, then injects the result with noise sampled from a Laplace distribution to produce a differentially private result. The system iteratively calibrates the noise value of the differentially private result using a secondary distribution different from the Laplace distribution and a new fractional privacy spend. The system ceases to iterate when an iteration uses the maximum privacy spend or a relative error of the differentially private result is determined to satisfy the target accuracy, or both. The system sends the differentially private result to the client.

ABSTRACT

A differentially private security system communicatively coupled to a database storing restricted data receives a database query from a client. The database query includes an operation, a target accuracy, and a maximum privacy spend for the query. The system performs the operation to produce a result, then injects the result with noise sampled from a Laplace distribution to produce a differentially private result. The system iteratively calibrates the noise value of the differentially private result using a secondary distribution different from the Laplace distribution and a new fractional privacy spend. The system ceases to iterate when an iteration uses the maximum privacy spend or a relative error of the differentially private result is determined to satisfy the target accuracy, or both. The system sends the differentially private result to the client.

Date Reçue/Date Received 2021-02-11

ADAPTIVE DIFFERENTIALLY PRIVATE COUNT

[0001]

BACKGROUND

FIELD OF DISCLOSURE

[0002] The present invention generally relates to computer database security and in particular to increasing differentially private database performance by bounding database query privacy spend.

DESCRIPTION OF THE RELATED ART

[0003] Data about people, such as health data, financial records, location information, web browsing, and viewing habits, is valuable for analysis and collaboration. There are many technologies in which statistical or predictive analysis of personal data is beneficial. For example, medical research institutions use medical information about populations of individuals to support epidemiologic studies. Map providers use location information gathered from mobile devices carried by people to determine traffic information and provide routing guidance. Technology companies collect information describing behaviors of Internet users to improve their offerings, such as by redesigning user interfaces to improve human-computer interactions, making improved recommendations, and offering sponsored messages.

Date Reçue/Date Received 2022-05-30

[0004] However, the personal nature of this data limits its usefulness. Government regulations provide strict rules about how personal data can be collected, used, and shared. Individuals also have expectations about how their personal data will be used, and may react negatively if it is publicly disclosed. As a result, companies that collect and maintain personal data seek ways to extract value from it without running afoul of such rules and expectations.

[0005] One set of techniques for using personal data involves removing personally-identifiable information from the data through masking, hashing, anonymization, aggregation, and tokenization. These techniques tend to be resource intensive and may compromise analytical utility. For example, data masking may remove or distort data, compromising the statistical properties of the data. These techniques also often fail to protect individual privacy.

[0006] An additional technique makes use of differential privacy. Differential privacy is technology that injects noise into results provided by statistical databases in order to protect private information. Within this technological space, issues arise over how to evaluate the privacy impact of the injected noise. The answer can be complex due to the potential resources available to determined adversaries (e.g., the computing power available to a potential attacker trying to gain access to the private data), the resources (e.g., computing power) available to the database, and the types of queries supported by the database.

[0007] A differentially private system provides differentially private results in response to database queries. The amount of private information provided by the system may depend, in part, on a "privacy budget" that describes an amount of privacy that may be "spent" to retrieve information from the database. It is important for the differentially private system to calculate privacy spend correctly because it directly impacts the analytical utility of the information in the

Date Reçue/Date Received 2021-02-11

database. It is likewise important to for the system to minimize privacy spend to the extent possible in order to provide privacy budget for additional queries for the same reason.

SUMMARY

[0008] A differentially private security system communicatively coupled to a database storing restricted data receives a database query from a client. The database query includes an operation, a target accuracy, and a maximum privacy spend for the query. The system performs the operation to produce a result, then injects the result with noise sampled from a Laplace distribution to produce a differentially private result. The system iteratively calibrates the noise value of the differentially private result using a secondary distribution different from the Laplace distribution and a new fractional privacy spend. The system ceases to iterate when an iteration uses the maximum privacy spend or a relative error of the differentially private result is determined to satisfy the target accuracy, or both. The system sends the differentially private result to the client.

[0009] Calibrating the noise value of the differentially private result using the secondary distribution different from the Laplace distribution and the new fractional privacy spend larger than the one or more fractional privacy spends of one or more earlier iterations involves the system generating the new fractional privacy spend such that it is larger than any fractional privacy spends of preceding iterations. The system generates a new noise value sampled from the secondary distribution using the new fractional privacy spend. The system incorporates the new noise value into the differentially private result. The system then checks whether the calibrated differentially private result satisfies the target accuracy. Checking whether the calibrated differentially private result satisfies the target accuracy involves determining a relative error of

Date Reçue/Date Received 2021-02-11

the differentially private result using an error estimator and then determining whether the relative error is, at most, the target accuracy.

BRIEF DESCRIPTION OF DRAWINGS

[0010] FIG. 1 illustrates a system for receiving a query for a database and responding to the query by executing the query in a differentially private manner, according to one embodiment.

[0011] FIG. 2 illustrates an example database structure, according to one embodiment.

[0012] FIG. 3 illustrates an adaptive engine, according to one embodiment.

[0013] FIG. 4 illustrates a process for executing a query with adaptive differential privacy, according to one embodiment.

[0014] FIG. 5 is a block diagram illustrating components of an example machine able to read instructions from a machine readable medium and execute them in a processor or controller, according to one embodiment.

[0015] The figures depict embodiments of the invention for purposes of illustration only. One skilled in the art will readily recognize from the following description that alternative embodiments of the structures and methods illustrated herein may be employed without departing from the principles of the invention described herein.

DETAILED DESCRIPTION

[0016] Reference will now be made in detail to several embodiments, examples of which are illustrated in the accompanying figures. It is noted that wherever practicable similar or like reference numbers may be used in the figures and may indicate similar or like functionality.

Date Reçue/Date Received 2021-02-11

SYSTEM OVERVIEW

[0017] FIG. 1 is a system 100 for receiving a query 108 for a database 106 and responding to the query 108 by executing the query in a differentially private (DP) manner, according to one embodiment. The system 100 includes a differentially private security system (DP system) 102 that receives an analytical query 108 from a client 104 and applies a DP version of the query 114 on the database 106. Subsequently, the DP system 102 returns the response of the DP query 114 to the client 104 as the DP response 112.

[0018] The database 106 is one or more databases managed by one or more entities. The database 106 may be managed by the same entity that manages the DP system 102 or by a different entity. The database 106 stores at least some restricted data. The restricted data may be represented as rows of records, with each record having a set of columns holding values pertaining to the record.

[0019] Restricted data is data to which access and/or usage is limited due to legal, contractual, and/or societal concerns. Examples of restricted data include health data of patients and financial records of people, businesses or other entities. Similarly, restricted data may include census data or other forms of demographic data describing people, businesses, or other entities within geographic areas. Restricted data also includes usage data describing how people interact with electronic devices and/or network-based services. For example, restricted data may include location data describing geographic movements of mobile devices, consumption history data describing how and when people consume network-based content, and the particular content consumed (e.g., music and/or video content), and messaging data describing when and to whom users send messages via mobile or other electronic devices.

Date Reçue/Date Received 2021-02-11

[0020] A client 104 is used to access the restricted data in the database 106. A client 104 is an electronic device such as a desktop, laptop, or tablet computer or a smartphone used by a human user to access the database 106. The client 104 and user may be, but are not necessarily, associated with the entities that manage the database 106 and/or DP system 102. Users of the client 104 include administrators and analysts. Administrators use the clients 104 to access the DP system 102 and/or database 106 to perform administrative functions such as provisioning other users and/or clients 104, and configuring, maintaining, and auditing usage of the system and/or database. The administrators may access the DP system 102 and database 106 directly via administrative interfaces that allow users with appropriate credentials and access rights to perform the administrative functions.

[0021] Analysts use the clients 104 to apply analytical queries 108 to the restricted data in the database 106. The clients 104 used by the analysts access the database 106 only through the DP system 102. Depending upon the embodiment, the analyst and/or client 104 may have an account provisioned by an administrator which grants the analyst or client certain rights to access the restricted data in the database 106.

[0022] The rights to the restricted data may be specified in terms of a privacy budget. The privacy budget describes limits on how much of the restricted data can be released. In one embodiment, the privacy budget is a numerical value representative of a number and/or type of remaining queries 108 available, or a degree of information which can released about data, e.g., data in a database or accessible by the DP system 102. The privacy budget may be specified in terms of a query, analyst, client 104, entity, globally, and/or time period. For example, the privacy budget may specify limits for an individual query, with each query having a separate budget. The privacy budget may also specify limits for an analyst or client, in which case the

Date Reçue/Date Received 2021-02-11

budget is calculated cumulatively across multiple queries from a client or analyst. For a privacy budget specified for an entity, such as an organization having multiple clients 104 and users, the privacy budget is calculated cumulatively across the multiple queries from clients and users associated with the entity. A global privacy budget, in turn, is calculated across all queries to the database, regardless of the source of the query. The privacy budget may also specify an applicable time period. For example, the privacy budget may specify that queries from particular clients may not exceed a specified budget within a given time period, and the budget may reset upon expiration of the time period. Depending upon the embodiment, client, as used herein, may alternatively or additionally refer to a user using the client to access the DP system 102, to a user account registered with the DP system 102, to a group of users or to a group of clients 104, and/or to another entity that is a source of queries.

[0023] As discussed above, a client 104 sends an analytical query 108 to the DP system 102 and also receives a differentially private response 112 to the query from the system. The queries 108 submitted by the client 104 may be simple queries, such as count queries that request the number of entries in the databases 106 that satisfy a condition specified by the client 104, or complicated queries, such as predictive analytics queries that request a data analytics model trained on the databases 106. Specific types of queries are discussed in more detail below.

[0024] Each query has an associated set of privacy parameters. The privacy parameters indicate the amount of restricted data to release from the database 106 to the client 104 in response to the query 108. The privacy parameters likewise indicate a privacy spend, which is the amount of decrease in the relevant privacy budget (e.g., the budget for the client 104 or entity with which the client is associated) in response to performance of the query 108. In one embodiment, the client 104 specifies a set of associated privacy parameters with each submitted

Date Reçue/Date Received 2021-02-11

query 108. In other embodiments, the privacy parameters are specified in other ways. The DP system 102 may associate privacy parameters with received queries (rather than obtaining the parameters directly from the query). For example, the DP system 102 may apply a default set of privacy parameters to queries that do not specify the parameters. The values of the default privacy parameters may be determined based on the client 104, analyst, query type, and/or other factors, such as a privacy budget of the client.

[0025] The DP system 102 receives an analytical query 108 from the client 104 and returns a differentially private response 112 to the client. In one embodiment, the DP system 102 determines the privacy parameters associated with the query, and evaluates the parameters against the applicable privacy budget. Alternatively, the analytical query 108 may specify the one or more privacy parameters of the set of privacy parameters. If the analytical query 108 and associated privacy parameters exceeds the privacy budget, the DP system 102 may deny (i.e., not execute) the query. Alternatively, the DP system 102 may adjust the privacy parameters to fall within the privacy budget, and execute the query using the adjusted privacy parameters. If the privacy parameters do not exceed the privacy budget, the DP system 102 executes a DP version of the query 114 on the database 106, such that it releases a degree of restricted data from the database 106 indicated by the privacy parameters specified by the client 104, and also protects a degree of privacy of the restricted data specified by the privacy budget. For example, an administrator of the database 106 may set a privacy budget specifying a maximum threshold on the amount of restricted data released by given query 108 that the client 104 may not exceed. Thus, the DP system 102 balances privacy protection of the restricted data in the database 106 while releasing useful information on the database 106 to the client 104.

Date Reçue/Date Received 2021-02-11

[0026] The DP query 114 applied to the database 106 by the DP system 102 is a differentially private version of the query 108 that satisfies a definition of differential privacy described in more detail with reference to the privacy system 160 in FIG. 3. The DP system 102 may apply the DP query 114 to the database 106 by transforming the analytical query 108 into one or more queries derived from the analytical query that cause the database 106 to release differentially private results. The DP system 102 may then return these differentially private results to the client as the DP response 112. The DP system 102 may also, or instead, apply the DP query 114 to the database 106 by transforming the analytical query into one or more derived queries that cause the database to release results that are not necessarily differentially private. The DP system 102 may then transform the released results in a way that enforces differential privacy to produce the DP response 112 returned to the client 104. These transformations may involve perturbing the process by which the DP query 114 is produced from the analytical query 108 and/or perturbing the results released by the database 106 with noise that provides the differential privacy specified by the privacy parameters while enforcing the privacy budget.

[0027] The DP system 102 allows an analyst to perform database queries on restricted data, and thereby perform analyses using the DP responses 112 returned by the queries, while maintaining adherence with privacy parameters and a privacy budget. In addition, the techniques used by the DP system 102 allow database queries to access restricted data in ways that do not compromise the analytical utility of the data. The DP system 102 supports a wide variety of analytical and database access techniques and provides fine-grained control of the privacy parameters and privacy budget when using such techniques. The DP system 102 thus provides an improved database system having expanded and enhanced access to restricted data relative to other database systems.

Date Reçue/Date Received 2021-02-11

[0028] An analyst can use the DP system 102 for a variety of different purposes. In one embodiment, the restricted data in the database 106 includes training data describing features of entities relevant to a particular condition. The analyst uses the DP system 102 to build one or more differentially private machine-learned models, such as classifiers, from the training data. The analyst can apply data describing a new entity to the machine-learned models, and use the outputs of the models to classify the new entity as having, or not having the condition. However, an adversary cannot use the information in the machined-learned models to ascertain whether individual entities described by the training set have the condition due to the differentially private nature of the models.

[0029] Such models may be retained and executed within the DP system 102. For example, an analyst can issue an analytical query 108 that causes the DP system 102 to interact with the restricted data in the database 106 to build the machine-learned models. The DP system 102 can then store the models within the system or an associated system. The analyst can use a new analytical query 108 or another interface to the system 102 to apply the data describing the new entity to the models. The DP system 102 can execute the new data on the stored models and output the classification of the entity as a DP response 112. Alternatively or in addition, the DP system 102 can output the trained models as a DP response 112, and an analyst can store and apply data to the models using different systems in order to classify the entity.

[0030] Examples of the types of classifications that may be performed using such models include determining whether a person (the entity) has a medical condition. In this example, the restricted training data include health data describing patients that are labeled as having or not having a given medical condition. The analyst applies health data for a new patient to the one or

Date Reçue/Date Received 2021-02-11

more differentially private machine-learned models generated from the restricted training data in order to diagnose whether the new patient has the medical condition.

[0031] Another example classification that may be performed using such models involves identifying fraudulent or otherwise exceptional financial transactions. In this example, the restricted training data includes financial transaction data associated with one or more people or institutions, where the transactions are labeled as being exceptional or not exceptional. The analyst applies financial transaction data for a new transaction to the one or more differentially private machine-learned models generated from the restricted training data in order to determine whether the new transaction is exceptional. The analyst can block, flag, or otherwise report an exceptional transaction.

[0032] As shown in FIG. 1, the DP system 102 includes a user interface 150, a library 152, an account management system 154, a query handling engine 156, a data integration module 158, a privacy system 160, a count engine 162, and an adaptive engine 164. Some embodiments of the DP system 102 have different or additional modules than the ones described here. Similarly, the functions can be distributed among the modules in a different manner than is described here. Certain modules and functions can be incorporated into other modules of the DP system 102.

[0033] The user interface 150 generates a graphical user interface on a dedicated hardware device of the DP system 102 or the client 104 in which the client 104 can submit an analytical query 108 and the desired privacy parameters, view the DP response 112 in the form of numerical values or images, and/or perform other interactions with the system. The client 104 may also use the graphical user interface to inspect the database 106 schemata, view an associated privacy budget, cache the DP response 112 to view the response later, and/or perform

Date Reçue/Date Received 2021-02-11

administrative functions. The user interface 150 submits properly formatted query commands to other modules of the DP system 102.

[0034] The library 152 contains software components that can be included in external programs that allow the client 104 to submit the analytical query 108, receive the DP response 112, and other functions within a script or program. For example, the client 104 may use the software components of the library 152 to construct custom data analytic programs. Each of the software components in the library 152 submits properly formatted query commands to other modules of the DP system 102.

[0035] The account management system 154 receives properly formatted query commands (herein "query commands" or "QC"), parses the received query commands, and verifies that the commands are syntactically correct.

[0036] Examples of query commands accommodated by the DP system 102, according to one embodiment, are listed below.

QC1. Count

'SELECT COUNT (<column>) FROM <database.table> WHERE <where_clause> BUDGET <eps> <delta>.

QC2. Median

'SELECT MEDIAN (<column>) FROM <database.table> WHERE <where_clause> BUDGET <eps> <delta>.

QC3. Mean

'SELECT MEAN (<column>) FROM <database.table> WHERE <where_clause> BUDGET <eps> <delta>.

Date Reçue/Date Received 2021-02-11

QC4. Variance

'SELECT VARIANCE (<column>) FROM <database.table> WHERE <where_clause> BUDGET <eps> <delta>.

QC5. Inter-Quartile Range

'SELECT IQR (<column>) FROM <database.table> WHERE <where_clause> BUDGET <eps> <delta>.

QC6. Batch Gradient Descent

'SELECT <GLM> (<columns_x>,<column_y>,<params>) FROM <database.table> WHERE <where_clause> BUDGET <eps> <delta>.

QC7. Stochastic Gradient Descent

'SELECT SGD <GLM> (<column>) FROM <database.table> WHERE <where_clause> BUDGET <eps> <delta>.

QC8. Random Forest

'SELECT RANDOMFOREST (<columns_x>,<columns_y>) FROM <database.table> WHERE <where_clause> BUDGET <eps> <delta>.

QC9. Histogram

'SELECT HISTOGRAM (<column>) FROM <database.table> WHERE <where_clause_i> BUDGET <eps> <delta>.

[0037] The query handling engine 156 transforms the received query commands into appropriate function calls and database access commands by parsing the query command string. The function calls are specific to the query 108 requested by the client 104, and the access

Date Reçue/Date Received 2021-02-11

commands allow access to the required database 106. Different databases 106 require different access commands. The access commands are provided to the database integrator 158.

[0038] The database integrator 158 receives the access commands to one or more databases 106, collects the required databases, and merges them into a single data object. The data object has a structure similar to that of a database structure described in reference to FIG. 2. The data object is provided to the privacy system 160.

[0039] The privacy system 160 receives the data object from the database integrator 158, appropriate function calls from the query handling engine 156 indicating the type of query 108 submitted by the client 104, and privacy parameters specified for the query 108. The privacy system 160 evaluates the privacy parameters against the applicable privacy budget and either denies or allows the query. If the query is denied, the privacy system 160 outputs a response indicating that the query did not execute. If the query is allowed, the privacy system 160 executes the query and outputs a DP response 112 to a differentially private version of the query 108 with respect to the database 106. The privacy system 160 also decrements the applicable privacy budget to account for the executed query. The privacy system 160 uses differential privacy engines in the DP System 102, such as the count engine 162 and/or the adaptive engine 164, to execute the query. In an embodiment, the count engine 162 and/or adaptive engine 164 are components of the privacy system 160.

[0040] The count engine 162 generates a differentially private result in response to a query to count a set of data in the database 106, as described in greater detail below.

[0041] The adaptive engine 164 executes a query such that the DP system 102 pursues a target accuracy for results of the query. A target accuracy is specified in terms of a relative error. The

Date Reçue/Date Received 2021-02-11

target accuracy for a query is met if the differentially private result of the query has a relative error less than or equal to the target accuracy.

[0042] Relative error is the discrepancy between an exact value and an approximation of the exact value, in terms of a percentage. Specifically, relative error is:

$$\rho = \left|\frac{v_E - v_A}{v_E}\right| * 100\%$$

Where ρ is the relative error, v_E is the exact value, and v_A is the approximation. For example, assume a database stores information about patients in a hospital. A count query executed on the database requests a count of all patients in the hospital named Charles. The actual number of patients named Charles may be 100, but the DP system 102 provides a differentially private result with a value of 90. Here, $v_E = 100$ and $v_A = 90$. As such, the relative error ρ is 10%. This indicates that the differentially private result, 90, is 10% off from the exact value, 100.

[0043] A query executed by the adaptive engine 164 is an adaptive query that specifies a maximum privacy spend in terms of one or more privacy parameters, such as ε as described below, and a target accuracy in terms of a relative error percentage. For example, an adaptive query may specify a maximum privacy spend of $\varepsilon = 1$ and a target accuracy of 10%. The adaptive query also specifies one or more operations to perform on data and one or more relations indicating the data on which the adaptive engine 164 is to perform the one or more operations.

[0044] The adaptive engine 164 performs the operations and iteratively adjusts the noise added to the results, then checks whether the adjusted results of the operations satisfy the target accuracy. Each iteration uses a fraction of the maximum privacy spend. If the results of the operations at a given iteration do not satisfy the target accuracy, the adaptive engine 164 performs another iteration using a larger portion of the maximum privacy spend. The adaptive

Date Reçue/Date Received 2021-02-11

engine 164 ceases iterating when either the maximum privacy spend is spent or the target accuracy is achieved. For example, after a first iteration, 1/100 of the maximum privacy spend has been used and the results have a relative error of 20%, greater than a target accuracy of 10% relative error. As such, the adaptive engine 164 performs an additional iteration, spending 1/50 the maximum privacy spend. If the results of this second iteration have a relative error of 9%, the adaptive engine 164 ceases to iterate and provides the results of the second iteration to the client 104, as their relative error is within the target accuracy of 10%.

[0045] Using the techniques described herein, the DP system 102 can provide differentially private results that satisfy a target accuracy while minimizing the privacy spend. As such, the DP system 102 can avoid providing results that lack analytical utility due to a high amount of noise injected into the results. Simultaneously, the DP system 102 can avoid overspending privacy parameters to produce results for a query.

[0046] FIG. 2 illustrates an example database structure, according to one embodiment. The database 200 includes a data table, which may be referred to as a matrix, with a number of rows and columns. Each row is an entry of the database and each column is a feature of the database. Thus, each row contains a data entry characterized by a series of feature values for the data entry. For example, as shown in FIG. 2, the example database 200 contains a data table with 8 entries and 11 features, and illustrates a list of patient profiles. Each patient is characterized by a series of feature values that contain information on the patient's height (Feature 1), country of residence (Feature 2), age (Feature 10), and whether the patient has contracted a disease (Feature 11). A row is also referred to as a "record" in the database 106. The database 106 may include more than one data table. Henceforth a data table may be referred to as a "table."

Date Reçue/Date Received 2021-02-11

[0047] The feature values in the database 200 may be numerical in nature, e.g., Features 1 and 10, or categorical in nature, e.g., Features 2 and 11. In the case of categorical feature values, each category may be denoted as an integer. For example, in Feature 11 of FIG. 2, "0" indicates that the patient has not contracted a disease, and "1" indicates that the patient has contracted a disease.

DEFINITION OF DIFFERENTIAL PRIVACY

[0048] For a given query 108, the privacy system 160 receives a data object $\boldsymbol{X}$, function calls indicating the type of query 108, privacy parameters specified by the client 104, and outputs a DP response 112 to a differentially private version of the query 108 with respect to $\boldsymbol{X}$. Each data object $\boldsymbol{X}$ is a collection of row vectors $\boldsymbol{x}_{i=1, 2, \ldots, n}$, in which each row vector $\boldsymbol{x}_i$ has a series of p elements $x_i^{j=1,2,\ldots,p}$.

[0049] A query M satisfies the definition of ε-differential privacy if for all:

$$\forall \boldsymbol{X}, \boldsymbol{X}' \in \mathbb{D}, \forall S \subseteq Range(M): \frac{Pr[M(\boldsymbol{X}) \in S]}{Pr[M(\boldsymbol{X}') \in S]} \leq e^{\varepsilon}$$

where $\mathbb{D}$ is the space of all possible data objects, S is an output space of query M, and neighboring databases are defined as two data objects $\boldsymbol{X}$, $\boldsymbol{X}'$ where one of $\boldsymbol{X}$, $\boldsymbol{X}'$ has all the same entries as the other, plus one additional entry. That is, given two neighboring data objects $\boldsymbol{X}$, $\boldsymbol{X}'$ in which one has an individual's data entry (the additional entry), and the other does not, there is no output of query M that an adversary can use to distinguish between $\boldsymbol{X}$, $\boldsymbol{X}'$. That is, an output of such a query M that is differentially private reveals little to no information about individual records in the data object $\boldsymbol{X}$. The privacy parameter ε controls the amount of information that the query M reveals about any individual data entry in $\boldsymbol{X}$, and represents the degree of information released about the entries in $\boldsymbol{X}$. For example, in the definition given above, a small value of ε

Date Reçue/Date Received 2021-02-11

indicates that the probability an output of query M will disclose information on a specific data entry is small, while a large value of ε indicates the opposite.

[0050] As another definition of differential privacy, a query M is (ε,δ)-differentially private if for neighboring data objects $\boldsymbol{X}$, $\boldsymbol{X}'$:

$$\forall \boldsymbol{X}, \boldsymbol{X}' \in \mathbb{D}, \forall S \subseteq Range(M): \frac{Pr[M(\boldsymbol{X}) \in S]}{Pr[M(\boldsymbol{X}') \in S]} \leq e^{\varepsilon} + \delta.$$

The privacy parameter δ measures the improbability of the output of query M satisfying ε-differential privacy. As discussed in reference to FIG. 1, the client 104 may specify the desired values for the privacy parameters (ε,δ) for a query 108.

[0051] There are three important definitions for discussing the privacy system 160: global sensitivity, local sensitivity, and smooth sensitivity. Global sensitivity of a query M is defined as

$$GS_M(\boldsymbol{X}) = \max_{\boldsymbol{X},\boldsymbol{X}':d(\boldsymbol{X},\boldsymbol{X}')=1} \|M(\boldsymbol{X}) - M(\boldsymbol{X}')\|$$

where $\boldsymbol{X}$, $\boldsymbol{X}'$ are any neighboring data objects, such that $d(\boldsymbol{X}, \boldsymbol{X}')=1$. This states that the global sensitivity is the most the output of query M could change by computing M on $\boldsymbol{X}$ and $\boldsymbol{X}'$.

[0052] The local sensitivity of a query M on the data object $\boldsymbol{X}$ is given by:

$$LS_M(\boldsymbol{X}) = \max_{\boldsymbol{X}':d(\boldsymbol{X},\boldsymbol{X}')=1} \|M(\boldsymbol{X}) - M(\boldsymbol{X}')\|$$

where the set $\{\boldsymbol{X}': d(\boldsymbol{X}, \boldsymbol{X}')=1\}$ denotes all data objects that have at most one entry that is different from $\boldsymbol{X}$. That is, the local sensitivity $LS_M(\boldsymbol{X})$ is the sensitivity of the output of the query M on data objects $\boldsymbol{X}'$ that have at most one different entry from $\boldsymbol{X}$, measured by a norm function.

[0053] Related to the local sensitivity $LS_M(\boldsymbol{X})$, the smooth sensitivity given a parameter β is given by:

$$S_M(\boldsymbol{X};\beta) = \max_{\boldsymbol{X}'\in\mathbb{D}} \left\| LS_M(\boldsymbol{X}) \cdot e^{-\beta \cdot d(\boldsymbol{X},\boldsymbol{X}')} \right\|$$

Date Reçue/Date Received 2021-02-11

where $d(X, X')$ denotes the number of entries that differ between X and X'.

Notation for Random Variables

[0054] The notation in this section is used for the remainder of the application to denote the following random variables.

1) $G(\sigma^2)$, denotes a zero-centered Gaussian random variable with the probability density function

$$f(x|\sigma^2) = \frac{1}{\sigma\sqrt{2\,\pi}} e^{-\frac{x^2}{2\sigma^2}}.$$

2) $L(b)$ denotes a zero-centered Laplacian random variable from a Laplace distribution with the probability density function

$$f(x|b) = \frac{1}{2b} e^{-\frac{|x|}{b}}.$$

3) $C(\gamma)$ denotes a zero-centered Cauchy random variable with the probability density function

$$f(x|\gamma) = \frac{1}{\pi\gamma\left(1 + \left(\frac{x}{\gamma}\right)^2\right)}.$$

[0055] Further, a vector populated with random variables R as its elements is denoted by $\boldsymbol{v(R)}$. A matrix populated with random variables R as its elements is denoted by $M(R)$.

COUNT ENGINE

[0056] Turning back to FIG. 1, the count engine 162 produces a DP response 112 responsive to the differentially private security system 102 receiving a query 108 for counting the number of entries in a column of the data object $\boldsymbol{X}$ that satisfy a condition specified by the client 104, given privacy parameters ε and/or δ. An example query command for accessing the count engine 162 is given in QC1 above. For the example data object $\boldsymbol{X}$ shown in FIG. 2, the client 104 may

Date Reçue/Date Received 2021-02-11

submit a query 108 requesting a DP response 112 indicating the number of patients that are above the age of 30.

[0057] The count engine 162 retrieves the count q from $\boldsymbol{X}$. If privacy parameter δ is equal to zero or is not used, the count engine 162 returns

$$y \approx q + L(c_1 \cdot \frac{1}{\epsilon}),$$

as the DP response 112 for display by the user interface 150, where c_1 is a constant. An example value for c_1 may be 1. If the privacy parameter δ is non-zero, the count engine 302 returns

$$y \approx q + G\left(c_1 \cdot 2 \cdot log\frac{2}{\delta} \cdot \frac{1}{\epsilon^2}\right),$$

as the DP response 112 for display on the user interface 150, where c_1 is a constant. An example value for c_1 may be 1.

ADAPTIVE ENGINE

[0058] FIG. 3 illustrates an adaptive engine 164, according to one embodiment. The adaptive engine 164 includes an error estimator 310, an iterative noise calibrator 320, a secondary noise generator 330, and an accuracy manager 340. The adaptive engine 164 receives an adaptive query specifying a target accuracy in terms of a relative error value and a maximum privacy spend in terms of an ε value. The adaptive query also specifies a count operation to be performed on a set of data. Although described herein with reference to a count operation, the adaptive engine 164 can be used with alternative operations in alternative embodiments. Upon producing a differentially private result, the adaptive engine 164 sends the differentially private result to the client 104. The adaptive engine 164 may also send a notification identifying the relative error of the differentially private result.

Date Reçue/Date Received 2021-02-11

[0059] The error estimator 310 approximates the relative error of a differentially private result. Depending upon the embodiment, the error estimator 310 can be a plug-in estimator or a Bayesian estimator. The error estimator 310 generates a temporary result by applying the noise used to produce the differentially private result into the differentially private result. The error estimator 310 then determines a relative error between the differentially private result and the temporary result. The adaptive engine 164 uses this relative error to approximate the relative error of the differentially private result as compared to the original result.

[0060] The iterative noise calibrator 320 iteratively calibrates the noise of a differentially private result until the differentially private result has a relative error no greater than the target accuracy or the maximum privacy spend has been used, or both. Initially, the iterative noise calibrator 320 receives an initial differentially private result from a differentially private operation, such as a differentially private count performed by the count engine 162. The received initial differentially private result is broken down into its original result and the noise value injected into the original result to provide differential privacy. The iterative noise calibrator 320 also receives an indicator of a fraction of the maximum privacy spend which was used to generate the initial differentially private result. For example, the fraction of the maximum privacy spend, the "fractional privacy spend," may be 1/100 the maximum privacy spend S, i.e., $S/100$.

[0061] For a given iteration, the iterative noise calibrator 320 generates a corresponding fractional privacy spend such that it is larger than any fractional privacy spends of preceding iterations. For example, if the iterative noise calibrator 320 receives an indication that the fractional privacy spend to produce the initial differentially private result was $S/100$, a fractional privacy spend for a first iteration may be $S/50$, a fractional privacy spend for a second iteration

Date Reçue/Date Received 2021-02-11

may be $S/25$, and so on. The fractional privacy spend of an iteration increments by a specified amount from one iteration to the next. The increment can be based on the amount of the fractional privacy spend of an immediately preceding iteration. For example, the amount by which the fractional privacy spend of one iteration increases from a previous fractional privacy spend can be a doubling of the previous fractional privacy spend.

[0062] In an embodiment, the amount by which the fractional privacy spend of one iteration increases from a previous fractional privacy spend varies proportional to the difference between the target accuracy and a relative error of a differentially private result of a preceding iteration. The function by which the fractional privacy spend increases in proportion to the difference between the target accuracy and a relative error depends upon the embodiment. As an example of the variance, a first iteration produces a differentially private result with a relative error of 20%, where the target accuracy is 10%. As such, the fractional privacy spend may double. However, if after the first iteration the differentially private result has a relative error of 12%, then the second iteration may generate a fractional privacy spend that is only 20% larger than the fractional privacy spend used in the first iteration. In this second embodiment, the amount by which the fractional privacy spend can increase from one iteration to the next may be capped. For example, the fractional privacy spend may be capped to never more than double a preceding fractional privacy spend, e.g., $S/50$ will never be immediately followed by a larger fractional privacy spend than $S/25$, regardless of what the function outputs as the increment from the one fractional privacy spend to another.

[0063] For the given iteration, the iterative noise calibrator 320 generates a new noise value by sampling the secondary noise generator 330 using the new fractional privacy spend and the fractional privacy spend of the immediately preceding iteration (or, in the case of the first

Date Reçue/Date Received 2021-02-11

iteration, the fractional privacy spend indicated as used by the operation specified in the query). This sampling is described in greater detail below with reference to the secondary noise generator 330. The iterative noise calibrator 320 incorporates the new noise value into the differentially private result by injecting the new noise value into the original result from the operation specified in the query and updating the differentially private result to the resultant value.

[0064] After incorporating the new noise into the differentially private result, the iterative noise calibrator 320 checks whether the differentially private result satisfies the target accuracy using the error estimator 310. If the differentially private result satisfies the target accuracy by being no greater than the target accuracy, the iterative noise calibrator 320 ceases to iterate and sends the differentially private result to the client 104. If the differentially private result does not satisfy the target accuracy, the iterative noise calibrator 320 proceeds to another iteration.

[0065] If an iteration cannot increase the fractional privacy spend, i.e., the fractional privacy spend equals the maximum privacy spend, the iterative noise calibrator 320 stops iterating. If so, the adaptive engine 164 may send the differentially private result to the client 104 with a notification that the target accuracy could not be reached. The notification may indicate the achieved accuracy, i.e., the relative error.

[0066] The secondary noise generator 330 produces a secondary distribution different from the distribution used to produce the initial differentially private result. In an embodiment, the secondary distribution is a four-part mixture distribution. Specifically, the four-part mixture distribution may be one part Dirac delta function, two parts truncated exponential functions, and one part exponential function. In an embodiment, the distribution is as follows, where y is the

Date Reçue/Date Received 2021-02-11

new noise value, x is the previous noise value, a previous fractional privacy spend is $\varepsilon 1$, and a new fractional privacy spend is $\varepsilon 2$:

$$\frac{\varepsilon 1}{\varepsilon 2} e^{-(\varepsilon 2-\varepsilon 1)|x|}\delta(y-x) + \frac{\varepsilon 2^2 - \varepsilon 1^2}{2\varepsilon 2} e^{-\varepsilon 1|y-x|-\varepsilon 2|y|+\varepsilon 1|x|}$$

[0067] The iterative noise calibrator 320 samples the secondary noise generator 330 to generate the new noise value for injection into the result to provide differential privacy to the result. In an embodiment, the secondary noise generator 330 is sampled as follows, where a previous noise value is x, a new noise sample is y, a previous fractional privacy spend is $\varepsilon 1$, a new fractional privacy spend is $\varepsilon 2$, and z is drawn from the secondary distribution:

switch randomly
- **case** with probability $\frac{\varepsilon 1}{\varepsilon 2} e^{-(\varepsilon 2-\varepsilon 1)|x|}$:
  - return $y = x$.
- **case** with probability $\frac{\varepsilon 2-\varepsilon 1}{2\varepsilon 2}$:
  - draw $z \sim \{$
    - $e^{(\varepsilon 1+\varepsilon 2)z}$, for $z \leq 0$
    - 0, otherwise.
  - return $y = sgn(x)z$.
- **case** with probability $\frac{\varepsilon 1+ \varepsilon 2}{2\varepsilon 2}(1 - e^{-(\varepsilon 2- \varepsilon 1)|x|})$:
  - draw $z \sim \{$
    - $e^{-(\varepsilon 2-\varepsilon 1)z}$, for $0 \leq z \leq |x|$
    - 0, otherwise.
  - return $y = sgn(x)z$.
- **case** with probability $\frac{\varepsilon 2-\varepsilon 1}{2\varepsilon 2} e^{-(\varepsilon 2-\varepsilon 1)|x|}$:
  - draw $z \sim \{$
    - $e^{-(\varepsilon 1+\varepsilon 2)z}$, for $z \geq |x|$
    - 0, otherwise.
  - return $y = sgn(x)z$.

end switch

Date Reçue/Date Received 2021-02-11

PROCESSES

[0068] FIG. 4 illustrates a process for executing a query with adaptive differential privacy, according to one embodiment. The DP system 102 receives 410, from the client 104, a request to perform a query on a set of data. The query includes a target accuracy and a maximum privacy spend for the query. The DP system 102 performs 420 an operation to produce a result, such as a count operation, then injects the result with noise sampled from a Laplace distribution based on a fraction of the maximum privacy spend to produce a differentially private result.

[0069] The DP system 102 iteratively calibrates 430 the noise value of the differentially private result using a secondary distribution different from the Laplace distribution and a new fractional privacy spend. The new fractional privacy spend is generated to be larger than any fractional privacy spends of preceding iterations. The DP system 102 generates a new noise value sampled from the secondary distribution and incorporates it into the differentially private result to calibrate the noise of the differentially private result. The DP system 102 determines whether the calibrated differentially private result satisfies the target accuracy by determining a relative error of the calibrated differentially private result using an error estimator and comparing the relative error to the target accuracy. If the relative error is at most the target accuracy, the differentially private result satisfies the target accuracy.

[0070] The DP system 102 iterates until an iteration uses the maximum privacy spend or a relative error of the differentially private result is determined to satisfy the target accuracy, or both. The DP system 102 then sends 440 the differentially private result to the client 104 in response to the query. The DP system 102 may also send the relative error of the differentially private result to the client 104.

Date Reçue/Date Received 2021-02-11

COMPUTING ENVIRONMENT

[0071] FIG. 5 is a block diagram illustrating components of an example machine able to read instructions from a machine readable medium and execute them in a processor or controller, according to one embodiment. Specifically, FIG. 5 shows a diagrammatic representation of a machine in the example form of a computer system 500. The computer system 500 can be used to execute instructions 524 (e.g., program code or software) for causing the machine to perform any one or more of the methodologies (or processes) described herein. In alternative embodiments, the machine operates as a standalone device or a connected (e.g., networked) device that connects to other machines. In a networked deployment, the machine may operate in the capacity of a server machine or a client machine in a server-client network environment, or as a peer machine in a peer-to-peer (or distributed) network environment.

[0072] The machine may be a server computer, a client computer, a personal computer (PC), a tablet PC, a set-top box (STB), a smartphone, an internet of things (IoT) appliance, a network router, switch or bridge, or any machine capable of executing instructions 524 (sequential or otherwise) that specify actions to be taken by that machine. Further, while only a single machine is illustrated, the term "machine" shall also be taken to include any collection of machines that individually or jointly execute instructions 524 to perform any one or more of the methodologies discussed herein.

[0073] The example computer system 500 includes one or more processing units (generally processor 502). The processor 502 is, for example, a central processing unit (CPU), a graphics processing unit (GPU), a digital signal processor (DSP), a controller, a state machine, one or more application specific integrated circuits (ASICs), one or more radio-frequency integrated circuits (RFICs), or any combination of these. The computer system 500 also includes a main

Date Reçue/Date Received 2021-02-11

memory 504. The computer system may include a storage unit 516. The processor 502, memory 504 and the storage unit 516 communicate via a bus 508.

[0074] In addition, the computer system 506 can include a static memory 506, a display driver 510 (e.g., to drive a plasma display panel (PDP), a liquid crystal display (LCD), or a projector). The computer system 500 may also include alphanumeric input device 512 (e.g., a keyboard), a cursor control device 514 (e.g., a mouse, a trackball, a joystick, a motion sensor, or other pointing instrument), a signal generation device 518 (e.g., a speaker), and a network interface device 520, which also are configured to communicate via the bus 508.

[0075] The storage unit 516 includes a machine-readable medium 522 on which is stored instructions 524 (e.g., software) embodying any one or more of the methodologies or functions described herein. The instructions 524 may also reside, completely or at least partially, within the main memory 504 or within the processor 502 (e.g., within a processor's cache memory) during execution thereof by the computer system 500, the main memory 504 and the processor 502 also constituting machine-readable media. The instructions 524 may be transmitted or received over a network 526 via the network interface device 520.

[0076] While machine-readable medium 522 is shown in an example embodiment to be a single medium, the term "machine-readable medium" should be taken to include a single medium or multiple media (e.g., a centralized or distributed database, or associated caches and servers) able to store the instructions 524. The term "machine-readable medium" shall also be taken to include any medium that is capable of storing instructions 524 for execution by the machine and that cause the machine to perform any one or more of the methodologies disclosed herein. The term "machine-readable medium" includes, but not be limited to, data repositories in the form of solid-state memories, optical media, and magnetic media.

Date Reçue/Date Received 2021-02-11

CLAIMS

1. A method for performing an adaptive differentially private count operation on a set of data stored by a database, the method comprising:

receiving a request from a client device to perform a query on the set of data stored by the database, wherein the request identifies a target accuracy and a maximum privacy spend, wherein the target accuracy comprises a maximum relative error, and the maximum privacy spend comprises a value of a privacy parameter ε describing a degree of information released about the set of data due to the query;

performing, responsive to receiving the request to perform the query, a differentially private count operation on the set of data to produce a differentially private result, the differentially private count operation comprising:

performing a count operation on the set of data to produce a result;

perturbing the result to produce a differentially private result using a noise value sampled from a Laplace distribution and based on a fractional privacy spend comprising a fraction of the maximum privacy spend; and

iteratively calibrating the noise value of the differentially private result using a secondary distribution different from the Laplace distribution and a new fractional privacy spend until at least one of:

an iteration uses the maximum privacy spend, and

a relative error of the differentially private result is determined to satisfy the target accuracy; and

sending, to the client device, the differentially private result.

Date Reçue/Date Received 2022-05-30

2. The method of claim 1, wherein iteratively calibrating the noise value of the differentially private result using the secondary distribution different from the Laplace distribution comprises, for an iteration:

generating the new fractional privacy spend larger than the fractional privacy spends of preceding iterations;

generating a new noise value sampled from the secondary distribution using the new fractional privacy spend;

incorporating the new noise value into the differentially private result; and

determining whether the differentially private result satisfies the target accuracy.

3. The method of claim 2, wherein determining whether the differentially private result satisfies the target accuracy comprises:

estimating a relative error of the differentially private result; and

comparing the relative error to the target accuracy.

4. The method of claim 2, wherein the new fractional privacy spend is generated as a multiple of a preceding fractional privacy spend of a preceding iteration.

5. The method of claim 2, wherein the new fractional privacy spend is generated as a function of a difference between the target accuracy and a relative error of a differentially private result of a preceding iteration.

6. The method of claim 1, wherein the secondary distribution is a mixture distribution determined responsive to a plurality of functions.

Date Reçue/Date Received 2022-05-30

7. The method of claim 1, wherein iteratively calibrating the noise value of the differentially private result is responsive to a relative error of the differentially private result exceeding the target accuracy.

8. A non-transitory computer-readable medium storing computer-executable instructions that, when executed by one of more processors of a computing system, cause the computing system to perform operations for performing an adaptive differential private count operation on a set of data stored by a database, the operations comprising:

receiving a request from a client device to perform a query on the set of data stored by the database, wherein the request identifies a target accuracy and a maximum privacy spend, wherein the target accuracy comprises a maximum relative error, and the maximum privacy spend comprises a value of a privacy parameter ε describing a degree of information released about the set of data due to the query;

performing, responsive to receiving the request to perform the query, a differentially private count operation on the set of data to produce a differentially private result, the differentially private count operation comprising:

performing a count operation on the set of data to produce a result;

perturbing the result to produce a differentially private result using a noise value sampled from a Laplace distribution and based on a fractional privacy spend comprising a fraction of the maximum privacy spend; and

iteratively calibrating the noise value of the differentially private result using a secondary distribution different from the Laplace distribution and a new fractional privacy spend until at least one of:

an iteration uses the maximum privacy spend, and

Date Reçue/Date Received 2022-05-30

a relative error of the differentially private result is determined to satisfy the target accuracy; and

sending, to the client device, the differentially private result.

9. The non-transitory computer-readable medium of claim 8, wherein iteratively calibrating the noise value of the differentially private result using the secondary distribution different from the Laplace distribution comprises, for an iteration:

generating the new fractional privacy spend larger than the fractional privacy spends of preceding iterations;

generating a new noise value sampled from the secondary distribution using the new fractional privacy spend;

incorporating the new noise value into the differentially private result; and

determining whether the differentially private result satisfies the target accuracy.

10. The non-transitory computer-readable medium of claim 9, wherein determining whether the differentially private result satisfies the target accuracy comprises:

estimating a relative error of the differentially private result; and

comparing the relative error to the target accuracy.

11. The non-transitory computer-readable medium of claim 9, wherein the new fractional privacy spend is generated as a multiple of a preceding fractional privacy spend of a preceding iteration.

12. The non-transitory computer-readable medium of claim 9, wherein the new fractional privacy spend is generated as a function of a difference between the target accuracy and a relative error of a differentially private result of a preceding iteration.

Date Reçue/Date Received 2022-05-30

13. The non-transitory computer-readable medium of claim 8, wherein the secondary distribution is a mixture distribution determined responsive to a plurality of functions.

14. The non-transitory computer-readable medium of claim 8, wherein iteratively calibrating the noise value of the differentially private result is responsive to a relative error of the differentially private result exceeding the target accuracy.

15. A system, comprising:

a processor; and

a non-transitory computer-readable storage medium storing computer program instructions executable by a processor to perform operations for performing an adaptive differentially private count operation on a set of data stored by a database, the operations comprising:

receiving a request from a client device to perform a query on the set of data stored by the database, wherein the request identifies a target accuracy and a maximum privacy spend, wherein the target accuracy comprises a maximum relative error, and the maximum privacy spend comprises a value of a privacy parameter ε describing a degree of information released about the set of data due to the query;

performing, responsive to receiving the request to perform the query, a differentially private count operation on the set of data to produce a differentially private result, the differentially private count operation comprising:

performing a count operation on the set of data to produce a result;

perturbing the result to produce a differentially private result using a noise value sampled from a Laplace distribution and based on a fractional privacy spend comprising a fraction of the maximum privacy spend; and

Date Reçue/Date Received 2022-05-30

iteratively calibrating the noise value of the differentially private result using a secondary distribution different from the Laplace distribution and a new fractional privacy spend until at least one of:

an iteration uses the maximum privacy spend, and

a relative error of the differentially private result is determined to satisfy the target accuracy; and

sending, to the client device, the differentially private result.

16. The system of claim 15, wherein iteratively calibrating the noise value of the differentially private result using the secondary distribution different from the Laplace distribution comprises, for an iteration:

generating the new fractional privacy spend larger than the fractional privacy spends of preceding iterations;

generating a new noise value sampled from the secondary distribution using the new fractional privacy spend;

incorporating the new noise value into the differentially private result; and

determining whether the differentially private result satisfies the target accuracy.

17. The system of claim 16, wherein determining whether the differentially private result satisfies the target accuracy comprises:

estimating a relative error of the differentially private result; and

comparing the relative error to the target accuracy.

18. The system of claim 16, wherein the new fractional privacy spend is generated as a multiple of a preceding fractional privacy spend of a preceding iteration.

Date Reçue/Date Received 2022-05-30

19. The system of claim 16, wherein the new fractional privacy spend is generated as a function of a difference between the target accuracy and a relative error of a differentially private result of a preceding iteration.

20. The system of claim 15, wherein iteratively calibrating the noise value of the differentially private result is responsive to a relative error of the differentially private result exceeding the target accuracy.

Date Reçue/Date Received 2022-05-30

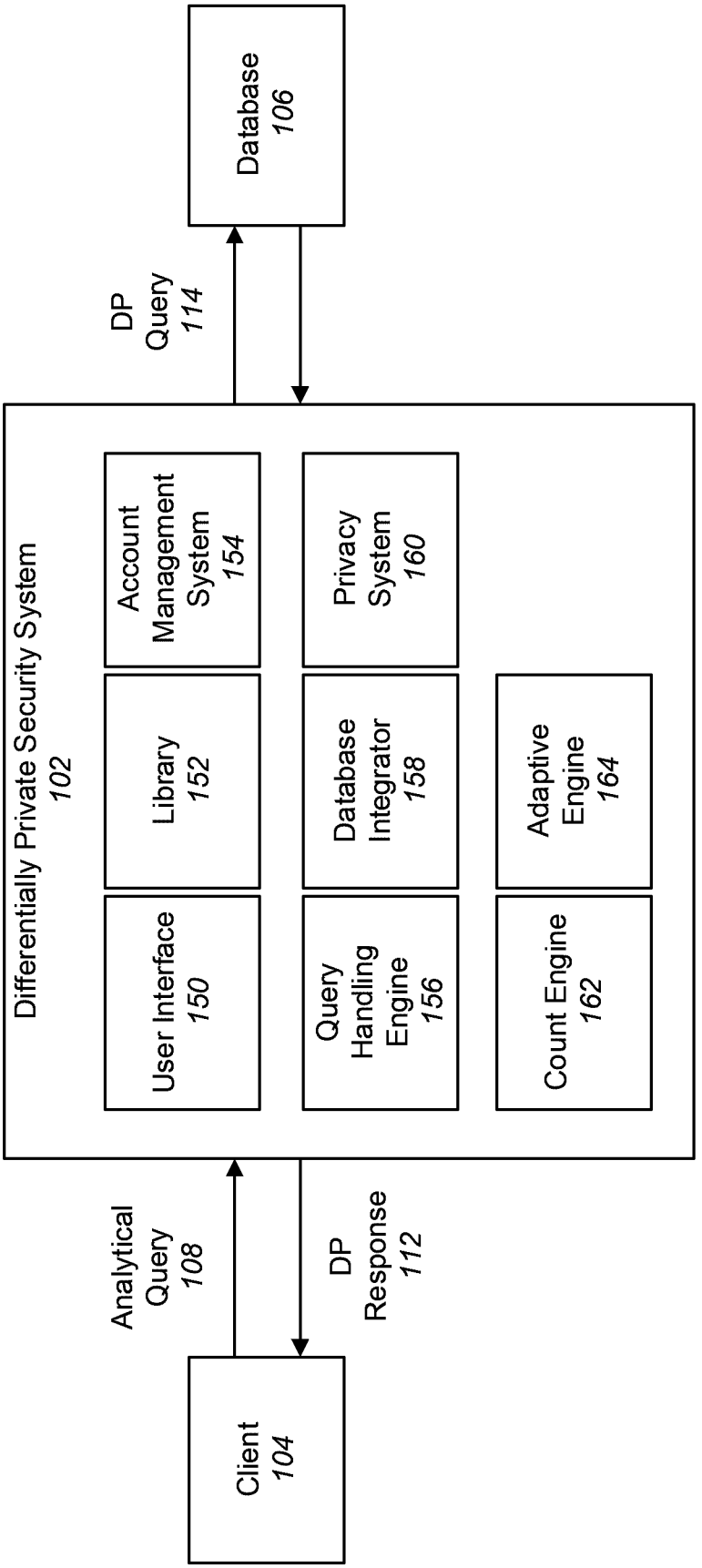


FIG. 1

Date Reçue/Date Received 2021-02-11

200

Entry #	Feature 1 (cm)	Feature 2 (residence)		Feature 10 (age)	Feature 11 (Disease)
1	163	Italy		37	0
2	136	England		87	0
3	180	France		54	1
4	347	USA	...	34	0
5	388	China		18	0
6	145	France		13	1
7	169	Korea		65	1
8	158	USA		17	1

FIG. 2

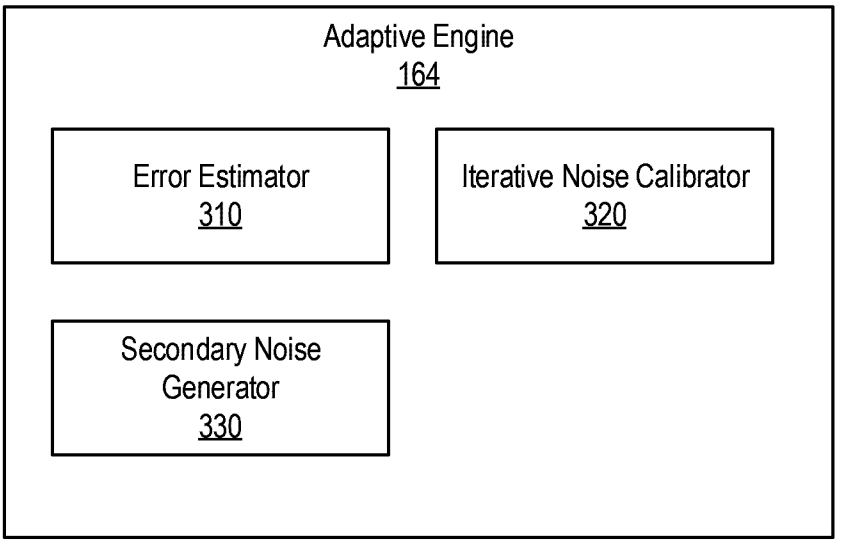


FIG. 3

Date Reçue/Date Received 2021-02-11

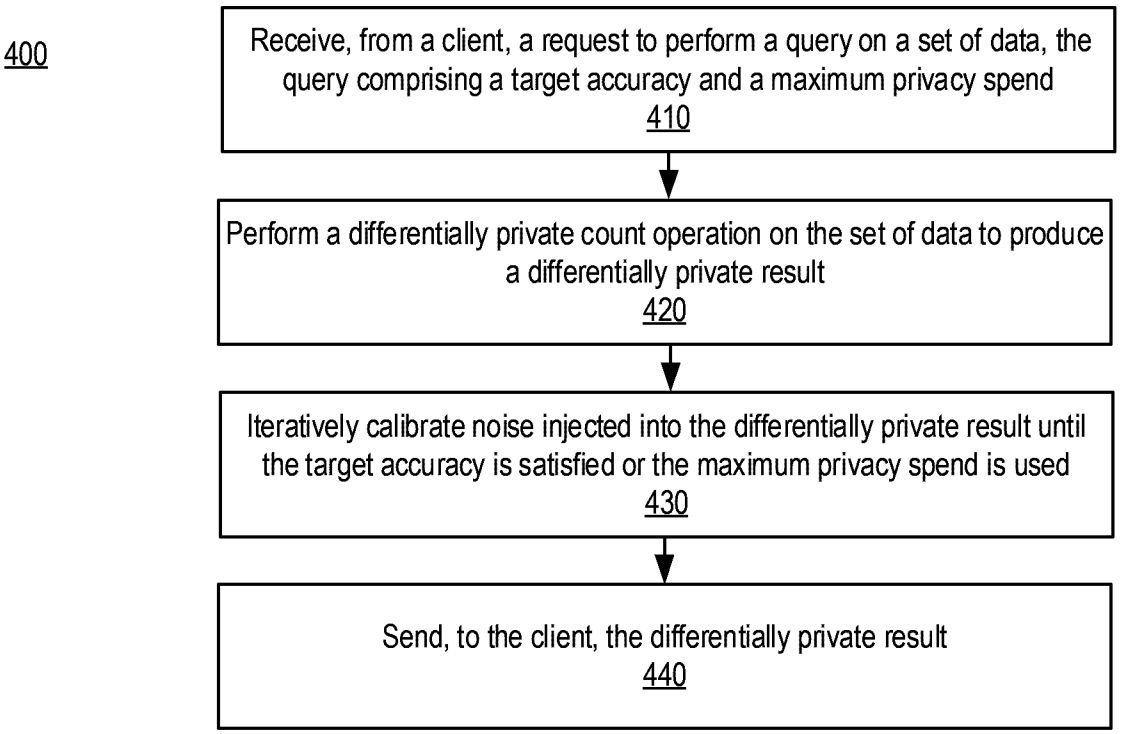


FIG. 4

Date Reçue/Date Received 2021-02-11

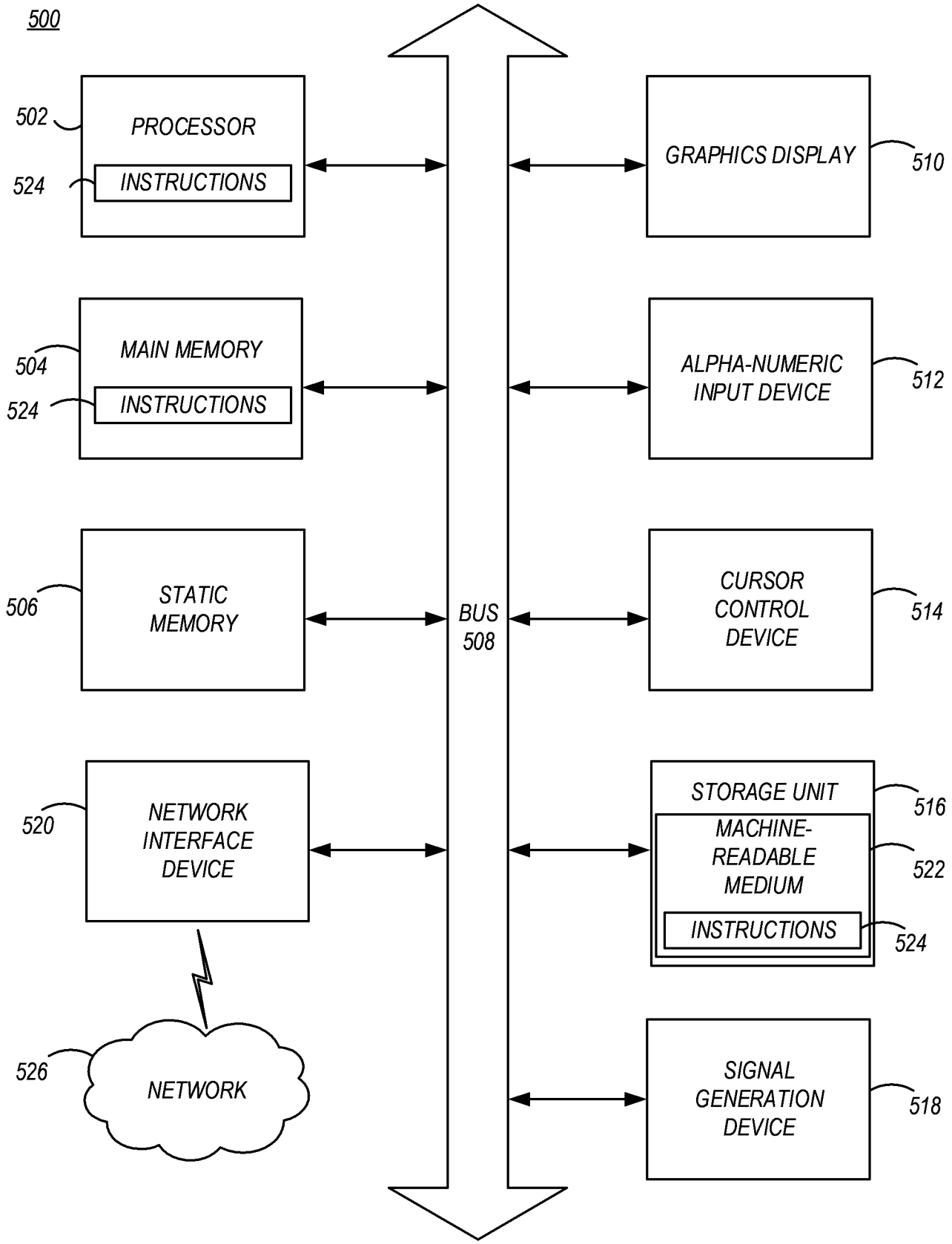
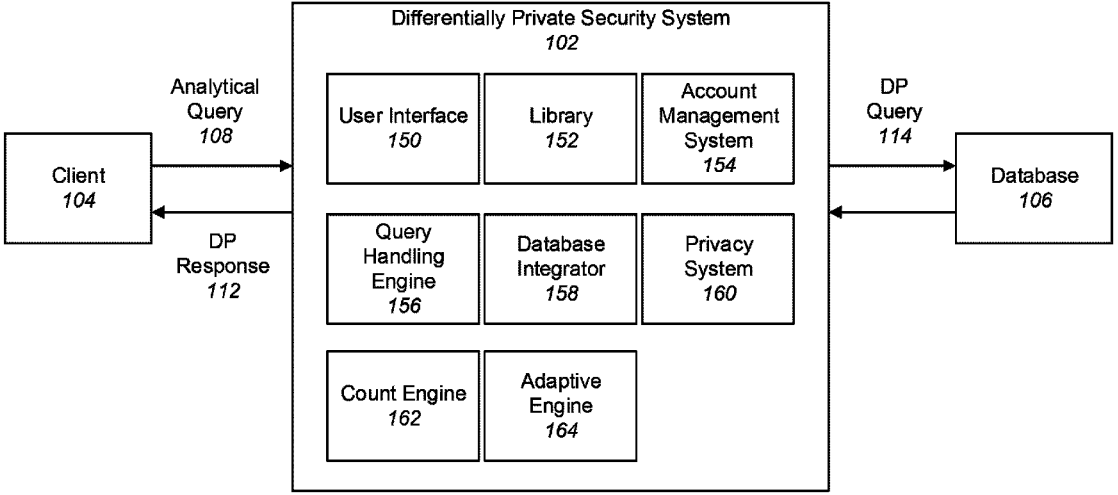


FIG. 5

Date Reçue/Date Received 2021-02-11


Differentially Private Security System
102
Client
104
Analytical
Query
108
DP
Response
112
User Interface
150
Library
152
Account
Management
System
154
Query
Handling
Engine
156
Database
Integrator
158
Privacy
System
160
Count Engine
162
Adaptive
Engine
164
DP
Query
114
Database
106
100