What is claimed is

1. A cost estimation method, comprising:

deciding whether acquired statistical information is complete, and in response to determining that the statistical information is incomplete, determining a cost estimation manner of a first operation type, wherein the first operation type is an operation type of an operation tree relying on the statistical information;

determining a first cost estimate of the first operation type based on the cost estimation manner; and

determining an accumulative cost estimate of the operation tree according to the first cost estimate and a second cost estimate of a second operation type, wherein the second operation type is an operation type of the operation tree not relying on the statistical information;

wherein the first operation type comprises a table scan operation, a filtering operation, a join operation and an aggregate operation; and the second operation type comprises a projection operation and a selection operation.

2. The method of claim 1, further comprising:

optimizing an execution plan corresponding to a structured query language (SQL) statement according to the accumulative cost estimate.

3. The method of claim 1, wherein determining the first cost estimate of the first operation type based on the cost estimation manner comprises:

determining a cost estimate of the table scan operation according to a record count in a data set;

determining a selection rate of a filtering condition according to a type of a filtering predicate, and determining a cost estimate of the filtering operation according to the selection rate;

determining a cost estimate of the join operation according to a determined record count of a join result set; and

determining an aggregate rate of aggregated fields according to the aggregated fields and an aggregate function, and determining a cost estimate of the aggregate operation according to the aggregate rate.

4. The method of claim 3, wherein determining the selection rate of the filtering condition

according to the type of the filtering predicate comprises:

deciding whether an object for calculating the selection rate corresponding to the type of the filtering predicate is available, and in response to determining that the object for calculating the selection rate is unavailable, determining the selection rate of the filtering condition as S = 1/a specified value.

5. The method of claim 4, wherein deciding whether the object for calculating the selection rate corresponding to the type of the filtering predicate is available, and in response to determining that the object for calculating the selection rate is unavailable, determining the selection rate of the filtering condition as S = 1/the specified value comprises:

in a case where the type of the filtering predicate is an identity predicate, deciding whether a number of distinct values of the filtering operation is available, and in response to determining that the number of distinct values of the filtering operation is unavailable, determining the selection rate of the filtering condition as S = 1/the specified value;

in a case where the type of the filtering predicate is a range predicate, in response to determining that an extreme value or a number of distinct values in a field of the filtering predicate is unavailable, determining the selection rate of the filtering condition as S = 1/the specified value; and

in a case where the type of the filtering predicate is a NULL predicate, deciding whether a number of null values in the field of the filtering predicate is available, and in response to determining that the number of null values in the field of the filtering predicate is unavailable, determining the selection rate of the filtering condition as S = 1/the specified value.

6. The method of claim 3, wherein determining the selection rate of the filtering condition according to the type of the filtering predicate comprises:

in response to the type of the filtering predicate being LIKE, determining the selection rate of the filtering condition as S = 1/a specified value;

in response to the type of the filtering predicate being an AND-cascaded predicate, determining the selection rate of the filtering condition according to a product of selection rates of cascaded predicates and a minimum value of filtering rates of the AND-cascaded predicate; and

in response to the type of the filtering predicate being an OR-cascaded predicate, determining the selection rate of the filtering condition according to selection rates of cascaded predicates corresponding to the OR-cascaded predicate.

7. The method of claim 3, before determining the cost estimate of the join operation according to the determined record count of the join result set, comprising:

in response to a number of distinct values in joined fields of a left table and a number of distinct values in joined fields of a right table joined to the left table being unavailable, determining the record count of the join result set based on a join type of the joined fields.

8. The method of claim 7, wherein determining the record count of the join result set based on the join type of the joined fields comprises:

in response to the join type of the joined fields being a join in a primary key-foreign key form, determining the record count of the join result set according to a record count of a foreign key field and a selection rate of a filtering condition of a primary key field;

in response to the join type of the joined fields being an inner join, using a larger value between a record count of the left table and a record count of the right table joined to the left table as the record count of the join result set;

in response to the join type of the joined fields being a cross join, using a product of the record count of the left table and the record count of the right table joined to the left table as the record count of the join result set;

in response to the join type of the joined fields being a left outer join, using the record count of the left table joined to the right table as the record count of the join result set, and in response to the join type of the joined fields being a right outer join, using the record count of the right table joined to the left table as the record count of the join result set; and

in response to the join type of the joined fields being a full join, using a sum of the record count of the left table and the record count of the right table joined to the left table as the record count of the join result set.

9. The method of claim 3, wherein determining the aggregate rate of the aggregated fields according to the aggregated fields and the aggregate function comprises:

in response to a set of aggregated fields comprising at least one primary key, determining the aggregate rate R_{agg} of the aggregated fields as $R_{agg} = 1$;

in response to the aggregate function corresponding to the aggregated fields being Group By, determining the aggregate rate R_{agg} of the aggregated fields as:

$$R_{agg} = \begin{cases} 1 & , n > 6 \\ 1 - \frac{1}{2^n} & , n \leqslant 6 \end{cases},$$

wherein n is a number of aggregated fields in the set of aggregated fields;

in response to the aggregate function corresponding to the aggregated fields being Rollup, determining the aggregate rate R_{agg} of the aggregated fields as:

$$R_{agg} = \sum_{k=1}^{n}\left(1 - \frac{1}{2^k}\right) = n - 1 + \frac{1}{2^n},$$

wherein n is the number of aggregated fields in the set of aggregated fields, and k is a positive integer; and

in response to the aggregate function corresponding to the aggregated fields being Cube, determining the aggregate rate R_{agg} of the aggregated fields as:

$$R_{agg} = \sum_{k=1}^{n} C_n^k \times \left(1 - \frac{1}{2^k}\right) = 2^n - \sum_{k=0}^{n} C_n^k \times \frac{1}{2^k},$$

wherein n is the number of aggregated fields in the set of aggregated fields, and k is a positive integer.

10. A cost estimation device, comprising:

a deciding means, which is configured to decide whether acquired statistical information is complete, and in response to determining that the statistical information is incomplete, determine a cost estimation manner of a first operation type, wherein the first operation type is an operation type of an operation tree relying on the statistical information;

a determining means, which is configured to determine a first cost estimate of the first operation type based on the cost estimation manner; and

an estimation means, which is configured to determine an accumulative cost estimate of the operation tree according to the first cost estimate and a second cost estimate of a second operation type, wherein the second operation type is an operation type of the operation tree not relying on the statistical information;

wherein the first operation type comprises a table scan operation, a filtering operation, a join operation and an aggregate operation; and the second operation type comprises a projection operation and a selection operation.

11. The device of claim 10, further comprising:

an execution means, which is configured to optimize an execution plan corresponding to a structured query language (SQL) statement according to the accumulative cost estimate.

12. The device of claim 10, wherein the determining means is configured to:

determine a cost estimate of the table scan operation according to a record count of a data set;

determine a selection rate of a filtering condition according to a type of a filtering predicate, and determine a cost estimate of the filtering operation according to the selection rate;

determine a cost estimate of the join operation according to a determined record count of a join result set; and

determine an aggregate rate of aggregated fields according to the aggregated fields and an aggregate function, and determine a cost estimate of the aggregate operation according to the aggregate rate.

13. The device of claim 12, wherein the determining means is configured to:

decide whether an object for calculating the selection rate corresponding to the type of the filtering predicate is available, and in response to determining that the object for calculating the selection rate is unavailable, determine the selection rate of the filtering condition as S = 1/a specified value.

14. The device of claim 13, wherein the determining means is configured to:

in a case where the type of the filtering predicate is an identity predicate, decide whether a number of distinct values of the filtering operation is available, and in response to determining that the number of distinct values of the filtering operation is unavailable, determine the selection rate of the filtering condition as S = 1/the specified value;

in a case where the type of the filtering predicate is a range predicate, in response to determining that an extreme value or a number of distinct values in a field of the filtering predicate is unavailable, determine the selection rate of the filtering condition as S = 1/the specified value; and

in a case where the type of the filtering predicate is a NULL predicate, decide whether a number of null values in the field of the filtering predicate is available, and in response to determining that the number of null values in the field of the filtering predicate is unavailable, determine the

selection rate of the filtering condition as S = 1/the specified value.

15. The device of claim 12, wherein the determining means is configured to:

in response to the type of the filtering predicate being LIKE, determine the selection rate of the filtering condition as S = 1/a specified value;

in response to the type of the filtering predicate being an AND-cascaded predicate, determine the selection rate of the filtering condition according to a product of selection rates of cascaded predicates and a minimum value of filtering rates of the AND-cascaded predicate; and

in response to the type of the filtering predicate being an OR-cascaded predicate, determine the selection rate of the filtering condition according to selection rates of cascaded predicates corresponding to the OR-cascaded predicate.

16. The device of claim 12, wherein the determining means is configured to:

in response to a number of distinct values in joined fields of a left table and a number of distinct values in joined fields of a right table joined to the left table being unavailable, determine the record count of the join result set based on a join type of the joined fields.

17. The device of claim 16, wherein the determining means is configured to:

in response to the join type of the joined fields being a join in a primary key-foreign key form, determine the record count of the join result set according to a record count of a foreign key field and a selection rate of a filtering condition of a primary key field;

in response to the join type of the joined fields being an inner join, use a larger value between a record count of the left table and a record count of the right table joined to the left table as the record count of the join result set;

in response to the join type of the joined fields being a cross join, use a product of the record count of the left table and the record count of the right table joined to the left table as the record count of the join result set;

in response to the join type of the joined fields being a left outer join, use the record count of the left table joined to the right table as the record count of the join result set, and in response to the join type of the joined fields being a right outer join, use the record count of the right table joined to the left table as the record count of the join result set; and

in response to the join type of the joined fields being a full join, use a sum of the record count of

the left table and the record count of the right table joined to the left table as the record count of the join result set.

18. The device of claim 12, wherein the determining means is configured to:

in response to a set of aggregated fields comprising at least one primary key, determine the aggregate rate R_{agg} of the aggregated fields as $R_{agg} = 1$;

in response to the aggregate function corresponding to the aggregated fields being Group By, determine the aggregate rate R_{agg} of the aggregated fields as:

$$R_{agg} = \begin{cases} 1 & , n > 6 \\ 1 - \frac{1}{2^n} & , n \leqslant 6 \end{cases},$$

wherein n is a number of aggregated fields in the set of aggregated fields;

in response to the aggregate function corresponding to the aggregated fields being Rollup, determine the aggregate rate R_{agg} of the aggregated fields as:

$$R_{agg} = \sum_{k=1}^{n}\left(1 - \frac{1}{2^k}\right) = n - 1 + \frac{1}{2^n},$$

wherein n is the number of aggregated fields in the set of aggregated fields, and k is a positive integer; and

in response to the aggregate function corresponding to the aggregated fields being Cube, determine the aggregate rate R_{agg} of the aggregated fields as:

$$R_{agg} = \sum_{k=1}^{n} C_n^k \times \left(1 - \frac{1}{2^k}\right) = 2^n - \sum_{k=0}^{n} C_n^k \times \frac{1}{2^k},$$

wherein n is the number of aggregated fields in the set of aggregated fields, and k is a positive integer.

19. A computing based device, comprising:

a processor, and

a memory, which is configured to store computer-executable instructions which, when executed, cause the processor to:

decide whether acquired statistical information is complete, and in response to determining that the statistical information is incomplete, determine a cost estimation manner of a first operation type, wherein the first operation type is an operation type of an operation tree relying on the statistical information;

determine a first cost estimate of the first operation type based on the cost estimation manner; and

determine an accumulative cost estimate of the operation tree according to the first cost estimate and a second cost estimate of a second operation type, wherein the second operation type is an operation type of the operation tree not relying on the statistical information;

wherein the first operation type comprises a table scan operation, a filtering operation, a join operation and an aggregate operation; and the second operation type comprises a projection operation and a selection operation.