

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 808 948**

51 Int. Cl.:

G06F 16/2453 (2009.01)

G06F 16/28 (2009.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **21.04.2016 PCT/CN2016/079812**

87 Fecha y número de publicación internacional: **10.11.2016 WO16177279**

96 Fecha de presentación y número de la solicitud europea: **21.04.2016 E 16789273 (6)**

97 Fecha y número de publicación de la concesión europea: **17.06.2020 EP 3293641**

54 Título: **Método y sistema de procesamiento de datos**

30 Prioridad:

04.05.2015 CN 201510222356

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

02.03.2021

73 Titular/es:

**ADVANCED NEW TECHNOLOGIES CO., LTD.
(100.0%)
Cayman Corporate Centre, 27 Hospital Road
George Town, Grand Cayman KY1-9008, KY**

72 Inventor/es:

HAN, MIN

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 808 948 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Método y sistema de procesamiento de datos

5 **Campo técnico**

La presente solicitud se refiere al campo de las tecnologías de macrodatos y, en particular, a un método y sistema de procesamiento de datos.

10 **Antecedentes de la técnica**

Con el desarrollo de tecnologías informáticas, el volumen de datos que necesita procesar una computadora es cada vez mayor, y una sola computadora no ha sido capaz de procesar datos a gran escala. Por lo tanto, se desarrolla una tecnología para construir una agrupación de computadoras combinando varias computadoras para procesar datos a gran escala en paralelo.

15 Una arquitectura de sistema de agrupación distribuida de Hadoop es una arquitectura de sistema de este tipo. Un sistema Hadoop puede construir una agrupación de computadoras utilizando una gran cantidad de computadoras baratas y reemplazar una computadora costosa de alta velocidad de computación con esta agrupación para realizar cálculos y almacenamiento de alta velocidad. El sistema Hadoop incluye principalmente un sistema de archivos distribuido y un sistema MapReduce. El sistema de archivos distribuido gestiona y almacena datos. El sistema MapReduce calcula los datos ingresados por el sistema de archivos distribuido, que incluye principalmente:
 20 descomponer un conjunto de datos a ser procesados en una pluralidad de bloques de datos; asignar cada una de las piezas de datos de pares de clave-valor originales en cada uno de los bloques de datos, para obtener datos de pares de clave-valor de resultado intermedio correspondientes a cada una de las piezas de datos de pares de clave-valor originales; y después de obtener los datos de pares de clave-valor de resultado intermedio correspondientes a todos los datos de pares de clave-valor original, reduciendo correspondientemente todos los datos de pares de clave-valor de resultado intermedio para obtener los correspondientes datos de pares de clave-valor de resultado final.

30 En la forma de procesamiento anterior, una gran tarea puede dividirse en una gran cantidad de tareas pequeñas y las tareas pequeñas se ejecutan por una gran cantidad de computadoras (también conocidas como ejecutores de tareas) en un sistema distribuido. De esta manera, se puede implementar el procesamiento rápido de datos masivos. Esta forma de procesamiento todavía no reduce los recursos de cálculo totales, sino que distribuye una gran cantidad de recursos de cálculo requeridos a una gran cantidad de computadoras, lo que comprime en gran medida el tiempo de procesamiento requerido. Esta forma de procesamiento es adecuada para un escenario fuera de línea insensible al tiempo. Para un escenario de servicio en línea, p. ej., un escenario de mensajería instantánea, generalmente se requiere que se realice el procesamiento masivo de datos y que se genere un resultado en poco tiempo; por lo tanto, es sensible al tiempo.

40 En el proceso de implementación de la presente solicitud, el inventor encuentra que la técnica anterior tiene al menos el siguiente problema:
 En un escenario de servicio en línea sensible al tiempo, una gran cantidad de recursos de cálculo todavía están ocupados para realizar el procesamiento de datos masivos, es decir, el volumen de datos procesados sigue siendo enorme. Por lo tanto, el proceso en el que el sistema Hadoop procesa datos consume mucho tiempo, de manera que un sistema de servicio tarda mucho tiempo en llamar al sistema Hadoop y esperar un resultado de procesamiento de datos, la eficiencia de ejecución es baja y el requisito específico de no se pueden cumplir los servicios fluidos, lo que resulta en una experiencia de usuario deficiente.

50 Por lo tanto, sobre la base de la investigación sobre el método de procesamiento de datos existente, el inventor proporciona un método y sistema de procesamiento de datos que tiene una alta eficiencia de ejecución y una experiencia de usuario deseable.

El documento US 2012/3041186 A1 describe técnicas para planificar uno o más trabajos de MapReduce en presencia de una o más clases de prioridad. Las técnicas incluyen la obtención de una clasificación preferida para uno o más trabajos de MapReduce, en donde la clasificación preferida comprende una o más clases de prioridad, priorizando la una o más clases de prioridad sujetas a una o más garantías de intervalo mínimo dinámico para cada una de las clases de prioridad, y empleando iterativamente un planificador de MapReduce, una vez por clase de prioridad, en orden de clase de prioridad, para optimizar el rendimiento del uno o más trabajos de MapReduce.

60 El documento JP 2010 092222 A describe una estructura de caché que puede utilizar efectivamente una caché en el proceso de Map-Reduce en base a la frecuencia de actualización y un método para construir un mecanismo de caché. El método incluye que una pluralidad de datos a ser procesados se agrupe en una pluralidad de grupos en base a la frecuencia de actualización de cada uno de los datos. El método incluye además un paso de calcular la frecuencia de actualización de grupo, que es la frecuencia de actualización de cada uno de la pluralidad de grupos, en base a la frecuencia de actualización de datos que es la frecuencia de actualización de los datos que constituyen cada uno de la pluralidad de grupos, generando un resultado parcial de una etapa de procesamiento de Map-Reduce para un grupo cuya frecuencia de actualización es igual o menor que un umbral y almacenar en caché el resultado parcial generado.

Resumen de la invención

5 Las realizaciones de la presente solicitud proporcionan un método y aparato de procesamiento de datos que tiene una alta eficiencia de ejecución y una experiencia de usuario deseable. El método y el aparato de procesamiento de datos están definidos por las reivindicaciones adjuntas.

El método y sistema de procesamiento de datos provisto en las realizaciones de la presente solicitud tiene al menos los siguientes efectos beneficiosos:

10 El sistema de procesamiento de datos procesa previamente los pares de clave-valor de punto caliente para facilitar la llamada por un sistema de servicio, mientras que los pares de clave-valor que no son de punto caliente se procesan solo cuando se llaman por el sistema de servicio, lo que reduce el volumen de datos que necesita procesarse en tiempo real mediante el sistema de procesamiento de datos que proporciona un servicio servidor para el sistema de servicio, mejora la eficiencia de ejecución del procesamiento de datos, reduce el tiempo del sistema de servicio esperando un resultado de procesamiento de datos y tiene un procesamiento de servicio fluido y una experiencia de usuario deseable.

Breve descripción de los dibujos

20 Los dibujos adjuntos descritos en el presente documento se utilizan para proporcionar una mejor comprensión de la presente solicitud, y construyen una parte de la presente solicitud. Las realizaciones ejemplares de la presente solicitud y las ilustraciones de la misma se utilizan para explicar la presente solicitud, y no pretenden presentar una limitación inadecuada a la presente solicitud. En los dibujos adjuntos:

25 la FIG. 1 es un diagrama de flujo de un método de procesamiento de datos de acuerdo con una realización de la presente solicitud;

la FIG. 2 es un diagrama de flujo de selección de una parte de los pares de clave-valor asignación como pares de clave-valor de punto caliente de acuerdo con una realización de la presente solicitud; y

30 la FIG. 3 es un diagrama estructural esquemático de un sistema de procesamiento de datos de acuerdo con una realización de la presente solicitud.

Descripción detallada

35 Para resolver los siguientes problemas técnicos en el método de procesamiento de datos existente: tiempo de procesamiento de datos largo, baja eficiencia de ejecución, incapacidad para cumplir con el requisito específico de servicios fluidos y experiencia de usuario deficiente, las realizaciones de la presente solicitud proporcionan un método de procesamiento de datos que procesa previamente los pares de clave-valor de punto caliente para facilitar la llamada mediante un sistema de servicio, mientras que los pares de clave-valor que no son de punto caliente se procesan solo cuando se llaman por el sistema de servicio, lo que reduce el volumen de datos que el sistema de procesamiento de datos necesita procesar en tiempo real al proporcionar un servicio servidor para el sistema de servicio, mejora la eficiencia de ejecución del procesamiento de datos, reduce el tiempo que el sistema de servicio esperando un resultado de procesamiento de datos y tiene un procesamiento de servicio fluido y una experiencia de usuario deseable.

45 Para hacer que los objetivos, las soluciones técnicas y las ventajas de la presente solicitud sean más comprensibles, a continuación, las soluciones técnicas de la presente solicitud se describen clara y completamente a través de realizaciones específicas de la presente solicitud y los correspondientes dibujos adjuntos. Aparentemente, las realizaciones descritas son simplemente algunas en lugar de todas las realizaciones de la presente solicitud.

Un sistema Hadoop puede incluir:

50 un cliente terminal JobClient configurado para enviar un trabajo de Map-Reduce;
un rastreador de trabajos JobTracker que es un proceso de Java y está configurado para coordinar la ejecución de todo el trabajo;

55 un rastreador de tareas TaskTracker que es un proceso Java y está configurado para ejecutar una tarea del trabajo; y
un sistema de archivos distribuidos de Hadoop (HDFS) configurado para compartir un archivo relacionado con el trabajo entre procesos.

Un proceso de trabajo del sistema Hadoop puede incluir:

1. Envío de tareas

60 El terminal cliente solicita un nuevo código de trabajo del rastreador de trabajos, crea una nueva instancia de trabajo y llama a una función submitJob.

2. Inicialización de tareas

Al recibir la llamada de la función submitJob, el rastreador de trabajos adquiere e inicializa una tarea. El rastreador de trabajos crea una tarea y asigna un código de tarea.

65 **3. Asignación de tareas**

El rastreador de trabajos asigna la tarea al rastreador de tareas.

4. Ejecución de tareas

Después de asignarse con una tarea, el rastreador de tareas comienza a ejecutar la tarea. Durante la asignación, el rastreador de tareas llama a una función de asignación para procesar la tarea, es decir, procesar pares de clave-valor originales para generar pares de clave-valor de resultado intermedio, y emite los pares de clave-valor de resultado intermedio de acuerdo con una secuencia de valores clave. Luego, el rastreador de tareas llama a una función de reducción para procesar los pares de clave-valor de resultado intermedio para generar pares de clave-valor de resultado final.

5. Finalización de tareas

Después de obtener un informe del rastreador de tareas que indica que todas las tareas se ejecutan correctamente, el rastreador de trabajos finaliza el trabajo.

La FIG. 1 es un diagrama de flujo de un método de procesamiento de datos de acuerdo con una realización de la presente solicitud, que incluye específicamente los siguientes pasos:

S100: Una parte de los pares de clave-valor a ser procesados se seleccionan como pares de clave-valor de punto caliente de acuerdo con una regla de detección.

Los datos se incorporan como un atributo y un valor numérico que describe las propiedades de los datos, es decir, un par de clave-valor comúnmente descrito. El par de clave-valor incluye un valor clave que representa un atributo y un valor clave que representa el contenido del atributo. El contenido del atributo incluye, pero no se limita a, una lista, un mapa hash, una cadena de caracteres, un valor numérico, un valor booleano, una matriz de lista ordenada, un valor nulo y similares. Por ejemplo, {"nombre": "Wang Xiao'er"} denota datos de una persona cuyo "nombre" es "Wang Xiao'er".

En una realización específica, el paso de seleccionar una parte de los pares de clave-valor a ser procesados como pares de clave-valor de punto caliente de acuerdo con una regla de detección incluye específicamente que: se pueden seleccionar aleatoriamente varios pares de clave-valor a ser procesados como los pares de clave-valor de punto caliente. De hecho, juzgar si un par de clave-valor a ser procesado es un par de clave-valor de punto caliente es un proceso complejo, especialmente cuando hay millones o incluso cientos de millones de pares de clave-valor a ser procesados. En la realización de la presente solicitud, el sistema de procesamiento de datos selecciona aleatoriamente varios pares de clave-valor a ser procesados como pares de clave-valor de punto caliente, simplificando así el proceso de juzgar si un par de clave-valor a ser procesado es un par de clave-valor de punto caliente y mejorando la eficiencia del procesamiento de datos del método.

Los pares de clave-valor de punto caliente se procesan previamente para llamarse mediante un sistema de servicio, mientras que los pares de clave-valor que no son de punto de punto caliente se procesan solo cuando se llaman por el sistema de servicio, lo que reduce el volumen de datos que necesitan procesarse en tiempo real tiempo mediante el sistema de procesamiento de datos que proporciona un servicio servidor para el sistema de servicio, mejora la eficiencia de ejecución del procesamiento de datos, reduce el tiempo del sistema de servicio esperando un resultado de procesamiento de datos y tiene un procesamiento de servicio fluido y una experiencia de usuario deseable.

En otra realización específica, haciendo referencia a la FIG. 2, el paso de seleccionar una parte de los pares de clave-valor a ser procesados como pares de clave-valor de punto caliente de acuerdo con una regla de detección incluye específicamente:

S101: Un primer número de pares de clave-valor a ser procesados se seleccionan aleatoriamente como pares de clave-valor candidatos.

S102: Se cuenta una frecuencia a la que se llama cada uno de los pares de clave-valor entre los pares de clave-valor candidatos.

S103: Los pares de clave-valor candidatos están dispuestos de acuerdo con las frecuencias.

S104: Se selecciona un segundo número de pares de clave-valor que tienen frecuencias de llamada máximas de los pares de clave-valor candidatos como pares de clave-valor de punto caliente.

El primer número es mayor que el segundo número.

En la realización de la presente solicitud, primero, el sistema de procesamiento de datos selecciona aleatoriamente un primer número de pares de clave-valor de asignación como pares de clave-valor candidatos. El primer número generalmente corresponde a un servicio específico. Un valor numérico del primer número puede ser un valor fijo establecido de acuerdo con la experiencia histórica, y también puede ser un valor numérico generado por una computadora a través de ajustes y cambios dinámicos.

Luego, el sistema de procesamiento de datos cuenta una frecuencia a la que se llama a cada uno de los pares de clave-valor entre los pares de clave-valor candidatos, y dispone los pares de clave-valor candidatos de acuerdo con las frecuencias. En una actividad de servicio específica, generalmente se requiere llamar a un número de pares de clave-valor para soportar el sistema de servicio. En este caso, el sistema de procesamiento de datos rastrea y registra la frecuencia con la que se llama a cada uno de los pares de clave-valor, es decir, el número de veces que se llama a

cada uno de los pares de clave-valor dentro de un período de tiempo. Además, el sistema de procesamiento de datos puede además disponer los pares de clave-valor de acuerdo con las frecuencias de llamada de mayor a menor.

A continuación, el sistema de procesamiento de datos selecciona un segundo número de pares de clave-valor que tienen frecuencias de llamada máximas de los pares de clave-valor candidatos como pares de clave-valor de punto caliente. El primer número es mayor que el segundo número. Del mismo modo, un valor numérico del segundo número puede ser un conjunto de valores fijos de acuerdo con la experiencia histórica, y también puede ser un valor numérico generado por el sistema de procesamiento de datos a través de ajustes y cambios dinámicos. Las frecuencias a las que se llaman a los pares de clave-valor de punto caliente seleccionados son mayores que las frecuencias a las que se llaman otros pares de clave-valor candidatos. El sistema de procesamiento de datos procesa previamente los pares de clave-valor de punto caliente, en lugar de procesar previamente otros pares de clave-valor candidatos. Una probabilidad de que se llamen a los pares de clave-valor procesados previamente es mayor que la probabilidad de que se llamen a otros pares de clave-valor. Por lo tanto, se reduce el volumen de datos que el sistema de procesamiento de datos necesita procesar en tiempo real al proporcionar un servicio servidor para el sistema de servicio, se mejora la eficiencia de ejecución del procesamiento de datos y se reduce el tiempo del sistema de servicio esperando un resultado de procesamiento de datos, el procesamiento del servicio es fluido y la experiencia del usuario es deseable.

Además, en la realización de la presente solicitud, el paso de seleccionar una parte de pares de clave-valor de asignación como pares de clave-valor de punto caliente incluye, además: establecer un conjunto de condiciones de categoría de servicio de pares de clave-valor candidatos antes del paso de seleccionar aleatoriamente un primer número de pares de clave-valor a ser procesados como pares de clave-valor candidatos; y seleccionar pares de clave-valor a ser procesados que cumplan con el conjunto de condiciones de categoría de servicio.

En la realización de la presente solicitud, el conjunto de condiciones de categoría de servicio puede ser un conjunto de valores fijos de acuerdo con la experiencia histórica, y también puede generarse a través de ajustes y cambios dinámicos. De hecho, un par de clave-valor llamado por un sistema de servicio de una actividad de servicio generalmente tiene algunas propiedades específicas para distinguirse de otras actividades de servicio. Por ejemplo, un par de clave-valor llamado por un sistema de servicio para enviar información tiene sus propiedades específicas en comparación con un par de clave-valor llamado por un sistema de servicio para el pago. El sistema de servicio para enviar información puede estar relacionado con un par de clave-valor que indica la edad de un receptor. Por ejemplo, la información enviada acerca de artículos de boda es generalmente información basura para receptores menores de 16 años. Cuando un conjunto de condiciones de categoría de servicio del sistema de servicio para enviar información incluye un par de clave-valor que indica la edad, se puede lograr un efecto de envío deseable.

Por lo tanto, se establece un conjunto de condiciones de categoría de servicio de pares de clave-valor a ser procesados, y el sistema de procesamiento de datos puede filtrar un gran número de pares de clave-valor a ser procesados mediante un juicio sobre el conjunto de condiciones de categoría de servicio, mejorando así la precisión de selección de pares de clave-valor de punto caliente. Por lo tanto, se reduce el volumen de datos que el sistema de procesamiento de datos necesita procesar en tiempo real al proporcionar un servicio servidor para el sistema de servicio, se mejora la eficiencia de ejecución del procesamiento de datos y se reduce el tiempo del sistema de servicio esperando un resultado de procesamiento de datos, el procesamiento del servicio es fluido y la experiencia del usuario es deseable.

En aún otra realización específica de la presente solicitud, el paso de seleccionar una parte de pares de clave-valor a ser procesados como pares de clave-valor de punto caliente de acuerdo con una regla de detección incluye específicamente:

establecer un umbral de frecuencia de llamada de los pares de clave-valor de punto caliente; y cuando la frecuencia de un par de clave-valor que se llama es mayor que el umbral de frecuencia de llamada, el par de clave-valor se establece como un par de clave-valor de punto caliente.

En la realización de la presente solicitud, se establece una frecuencia de llamada de pares de clave-valor, y cuando la frecuencia de un par de clave-valor que se llama es mayor que el umbral de frecuencia de llamada, el sistema de procesamiento de datos establece el par de clave-valor como un par de clave-valor de punto caliente. El sistema de procesamiento de datos procesa previamente los pares de clave-valor de punto caliente, en lugar de procesar previamente otros pares de clave-valor. Una probabilidad de que se llamen a los pares de clave-valor procesados previamente es mayor que la probabilidad de que se llamen a otros pares de clave-valor. Por lo tanto, se reduce el volumen de datos que el sistema de procesamiento de datos necesita procesar en tiempo real al proporcionar un servicio servidor para el sistema de servicio, se mejora la eficiencia de ejecución del procesamiento de datos y se reduce el tiempo del sistema de servicio esperando un resultado de procesamiento de datos, el procesamiento del servicio es fluido y la experiencia del usuario es deseable.

S200: Los pares de clave-valor de punto caliente se asignan para obtener pares de clave-valor de resultado intermedio correspondientes a los pares de clave-valor de punto caliente.

En una realización proporcionada en la presente solicitud, el terminal cliente JobClient envía un trabajo de Map-Reduce al rastreador de trabajos, crea una nueva instancia de trabajo y llama a una función submitJob. Al recibir la llamada

de la función submitJob, el rastreador de trabajos adquiere e inicializa una tarea. El rastreador de trabajos crea una tarea y asigna un código de tarea. El rastreador de trabajos asigna la tarea al rastreador de tareas. Después de asignarse con una tarea, el rastreador de tareas comienza a ejecutar la tarea. Durante la asignación, el rastreador de tareas llama a una función de asignación para procesar la tarea, es decir, procesar pares de clave-valor originales para generar pares de clave-valor de resultado intermedio, y emite los pares de clave-valor de resultado intermedio de acuerdo con una secuencia de valores clave.

S300: Los pares de clave-valor de resultado intermedio se reducen para generar pares de clave-valor de resultado final para la llamada.

En este paso, el rastreador de tareas llama a una función de reducción para procesar los pares de clave-valor de resultado intermedio para generar pares de clave-valor de resultado final. Después de obtener un informe del rastreador de tareas que indica que todas las tareas se ejecutan con éxito, el rastreador de trabajos almacena los pares de clave-valor de resultado final en el HDFS y finaliza el trabajo.

En la realización de la presente solicitud, el sistema de procesamiento de datos procesa previamente los pares de clave-valor de punto caliente para facilitar la llamada por un sistema de servicio, mientras que los pares de clave-valor que no son de punto caliente se procesan solo cuando se llaman por el sistema de servicio, lo que reduce el volumen de datos que el sistema de procesamiento de datos debe procesar en tiempo real al proporcionar un servicio servidor para el sistema de servicio, mejora la eficiencia de ejecución del procesamiento de datos, reduce el tiempo del sistema de servicio esperando un resultado de procesamiento de datos y tiene un procesamiento de servicio fluido y una experiencia de usuario deseable.

En una realización proporcionada en la presente solicitud, el método incluye, además: optimizar la regla de detección utilizando un modelo de aprendizaje automático.

El modelo de aprendizaje automático se relaciona con la inteligencia artificial. En la realización de la presente solicitud, la regla de detección se optimiza utilizando un modelo de aprendizaje automático. Después de que el sistema de procesamiento de datos se ejecute durante un período de tiempo, la precisión de juzgar los pares de clave-valor de punto caliente y los pares de clave-valor que no son de punto caliente puede mejorar significativamente. Por lo tanto, se reduce el volumen de datos que el sistema de procesamiento de datos necesita procesar en tiempo real al proporcionar un servicio servidor para el sistema de servicio, se mejora la eficiencia de ejecución del procesamiento de datos y se reduce el tiempo del sistema de servicio esperando un resultado de procesamiento de datos, el procesamiento del servicio es fluido y la experiencia del usuario es deseable.

El tipo de modelo de aprendizaje automático se selecciona de acuerdo con un sistema de servicio específico, y la optimización de un modelo de aprendizaje automático en la regla de detección se introduce de forma simple a continuación.

Específicamente, una condición de distribución de las frecuencias de los pares de clave-valor que se llaman frente a un solo atributo se cuenta utilizando un algoritmo de agrupación en el modelo de aprendizaje automático.

De acuerdo con la condición de distribución de las frecuencias de los pares de clave-valor que se llaman frente a un solo atributo, se selecciona un intervalo de los valores clave del contenido del atributo en el que las frecuencias de los pares de clave-valor que se llaman no son inferiores a un umbral de frecuencia predeterminado.

El intervalo de valores clave del contenido del atributo se establece como una condición de regla de la regla de detección.

La ilustración se realiza tomando todavía el sistema de servicio anterior para enviar información como ejemplo. Suponga que el sistema de servicio obtiene realizando estadísticas que los servicios para enviar información exceden una proporción preestablecida, p. ej., 50%, y se llama a un par de clave-valor que indica la edad de un receptor. El modelo de aprendizaje automático optimiza la regla de detección mediante un algoritmo de agrupación de K-medias.

Suponga que un conjunto de muestra (pares de clave-valor que indican las edades de los receptores y las frecuencias a las que se llaman a los pares de clave-valor) se clasifica en m categorías (segmentos de frecuencia), el algoritmo se describe de la siguiente manera:

(1) Los centros (frecuencias) iniciales de las m categorías (segmentos de frecuencia) se seleccionan correctamente.

(2) En la k-ésima iteración, se obtienen distancias (diferencias de frecuencia) desde cualquier muestra (un par de clave-valor que indica la edad de un receptor y una frecuencia a la que se llama al par de clave-valor) a los m centros, y la muestra (un par de clave-valor que indica la edad de un receptor y una frecuencia a la que se llama al par de clave-valor) se clasifica en una categoría (segmento de frecuencia) donde se encuentra un centro que tiene la distancia mínima.

(3) Un valor central (frecuencia) de la categoría (segmento de frecuencia) se actualiza utilizando un método de promedio.

(4) Para todos los m valores (frecuencias) centrales, si los valores de los mismos se mantienen sin cambios después de actualizarse utilizando el método de iteración de (2) y (3), la iteración termina; de lo contrario, la iteración continúa.

5 (5) Los centros (edades) iniciales de n categorías (grupos de edad) se seleccionan adecuadamente para cada una de las categorías (segmento de frecuencia) en m categorías (segmentos de frecuencia).

10 (6) En la k -ésima iteración, se obtienen distancias (diferencias de edad) desde cualquier muestra (un par de clave-valor que indica la edad de un receptor y una frecuencia a la que se llama al par de clave-valor) a los n centros, y la muestra (un par de clave-valor que indica la edad de un receptor y una frecuencia a la que se llama al par de clave-valor) se clasifica en una categoría (grupo de edad) donde se encuentra un centro que tiene la distancia mínima.

(7) Un valor (edad) central de la categoría (grupo de edad) se actualiza mediante un método de promedio.

15 (8) Para todos los n valores (edades) centrales, si los valores de los mismos se mantienen sin cambios después de actualizarse utilizando el método de iteración de (6) y (7), la iteración termina; de lo contrario, la iteración continúa.

20 Utilizando el algoritmo, se puede obtener a través del cálculo una regla de agrupación de pares de clave-valor a ser procesados que tienen frecuencias de llamada grandes frente a la edad. El hecho de que la edad del receptor se encuentre en una determinada categoría (grupo de edad) se utiliza como condición de la regla de detección. Por ejemplo, que la edad del receptor es 12-18 se utiliza como una condición de regla para juzgar que el par de clave-valor a ser procesado es un par de clave-valor de punto caliente. Después de optimizar la regla de detección utilizando el modelo de aprendizaje automático, el sistema de servicio detecta los pares de clave-valor de punto caliente a partir de los pares de clave-valor a ser procesados de acuerdo con la regla de detección optimizada.

25 En la realización proporcionada en la presente solicitud, un módulo de optimización de reglas está configurado además para:

30 cuando un par de clave-valor de un atributo y un par de clave-valor de otro atributo se llaman mediante los sistemas de servicio que tienen un mismo código de servicio, establecer una unión de conjuntos de intervalos de valores clave del contenido del atributo de los pares de clave-valor de los dos atributos como una condición de la regla de detección.

Suponga que el modelo de aprendizaje automático logra además una optimización en la regla de detección en la dimensión de las profesiones de los receptores después de lograr la optimización en la regla de detección en la dimensión de las edades de los receptores.

35 El sistema de procesamiento de datos calcula que un par de clave-valor a ser procesado que indica que el receptor se encuentra en un grupo de edad y un par de clave-valor a ser procesado que indica que el receptor es de una profesión están muy relacionados con el envío de información. Por ejemplo, un par de clave-valor a ser procesado indica que el receptor está en un grupo de edad de 20-30, un par de clave-valor a ser procesado indica que el receptor está en la industria de la informática, y cuando el sistema de servicio envía la información, se pueden lograr efectos deseables de promoción del servicio para receptores que tienen características de las dos dimensiones simultáneamente. Luego, el modelo de aprendizaje automático asocia el par de clave-valor a ser procesado que indica que el receptor está en el grupo de edad de 20-30 años con el par de clave-valor a ser procesado que indica que el receptor está en la industria de la informática, para formar un grupo de datos de pares de clave-valor de punto caliente.

45 El sistema de procesamiento de datos clasifica además las frecuencias a las que se llama a los grupos de datos de pares de clave-valor de punto caliente utilizando el modelo de aprendizaje automático, y clasifica los grupos de datos de pares de clave-valor de punto caliente en grupos de datos de punto caliente y grupos de datos que no son de punto caliente. Un modo de ajuste dinámico de los grupos de datos de punto caliente es: establecer un umbral de frecuencia de llamada de los grupos de datos de punto caliente, y cuando las frecuencias de los pares de clave-valor en el grupo de datos que se llama son mayores que el umbral de frecuencia, establecer el grupo de datos como un grupo de datos de punto caliente.

55 En la realización de la presente solicitud, se establece un valor de prioridad de procesamiento de un grupo de datos. El valor de prioridad se obtiene calculando un valor de suma ponderada de los pares de clave-valor a ser procesados. La prioridad de procesamiento del grupo de datos se ajusta dinámicamente de acuerdo con el valor de prioridad. Cuando un par de clave-valor en el grupo de datos se llama una vez, el valor de prioridad del grupo de datos se incrementa en una unidad. Cuando un valor de prioridad de un grupo de datos excede un valor de prioridad de un grupo de datos anterior del mismo, el sistema de procesamiento de datos mueve el grupo de datos una posición adelante. Mediante la optimización de la regla de detección utilizando el modelo de aprendizaje automático, los pares de clave-valor de punto caliente seleccionados mediante el sistema de procesamiento de datos a partir de los pares de clave-valor a ser procesados son de pares de clave-valor de asignación que tienen frecuencias máximas de ser llamados, en donde los grupos de datos de punto caliente formados asociando los pares de clave-valor son grupos de datos que tienen frecuencias máximas de ser llamados. Por lo tanto, se reduce el volumen de datos que el sistema de procesamiento de datos necesita procesar en tiempo real al proporcionar un servicio servidor para el sistema de servicio, se mejora la eficiencia de ejecución del procesamiento de datos y se reduce el tiempo del sistema de servicio

esperando un resultado de procesamiento de datos, el procesamiento del servicio es fluido y la experiencia del usuario es deseable.

5 En una realización proporcionada en la presente solicitud, el método incluye, además:
cuando se llama a un par de clave-valor que no es de punto caliente, se procesa el par de clave-valor que no es de punto caliente utilizando una función de reducción para generar datos para la llamada.

10 En la realización de la presente solicitud, el sistema de procesamiento de datos procesa previamente los pares de clave-valor de punto caliente utilizando la función de reducción, para generar datos para la llamada mediante el sistema de servicio. Cuando el sistema de servicio llama a los pares de clave-valor que no son de puntos caliente, el sistema de procesamiento de datos procesa los pares de clave-valor utilizando la función de reducción en tiempo real, para generar datos para la llamada mediante el sistema de servicio. Por lo tanto, se reduce el volumen de datos que el sistema de procesamiento de datos necesita procesar en tiempo real al proporcionar un servicio servidor para el sistema de servicio, se mejora la eficiencia de ejecución del procesamiento de datos y se reduce el tiempo del sistema de servicio esperando un resultado de procesamiento de datos, el procesamiento del servicio es fluido y la experiencia del usuario es deseable.

20 En una forma alternativa de la realización de la presente solicitud, un método de procesamiento de datos incluye los siguientes pasos:
asignar los pares de clave-valor a ser procesados para obtener pares de clave-valor de resultado intermedio correspondientes a los pares de clave-valor a ser procesados;
seleccionar una parte de los pares de clave-valor de resultado intermedio como pares de clave-valor de punto caliente de acuerdo con una regla de detección; y
25 reducir los pares de clave-valor de punto caliente para generar pares de clave-valor de resultado final para la llamada;
en donde el par de clave-valor incluye un valor clave que representa un atributo y un valor clave que representa el contenido del atributo.

30 Debe señalarse que una diferencia con respecto a la realización específica proporcionada anteriormente radica en que: el paso de seleccionar una parte de pares de clave-valor como los pares de clave-valor de punto caliente se establece para realizarse después del paso de procesamiento de asignación. En la realización de la presente solicitud, se reduce el volumen de datos de pares de clave-valor de procesamiento de reducción, y el problema del gran volumen puede resolverse hasta cierto punto. Por lo tanto, se reduce el volumen de datos que el sistema de procesamiento de datos necesita procesar en tiempo real al proporcionar un servicio servidor para el sistema de servicio, se mejora la eficiencia de ejecución del procesamiento de datos y se reduce el tiempo del sistema de servicio esperando un resultado de procesamiento de datos, el procesamiento del servicio es fluido y la experiencia del usuario es deseable.

35 El método de procesamiento de datos de acuerdo con las realizaciones de la presente solicitud se describe anteriormente. En base al mismo pensamiento, haciendo referencia a la FIG. 3, la presente solicitud proporciona además un sistema 1 de procesamiento de datos, que incluye:
40 un módulo 10 de detección configurado para seleccionar una parte de pares de clave-valor a ser procesados como pares de clave-valor de punto caliente de acuerdo con una regla de detección;
un módulo 20 de asignación configurado para asignar los pares de clave-valor de punto caliente para obtener pares de clave-valor de resultado intermedio correspondientes a los pares de clave-valor de punto caliente; y
45 un módulo 30 de reducción configurado para reducir los pares de clave-valor de resultado intermedio para generar pares de clave-valor de resultado final para la llamada;
en donde el par de clave-valor incluye un valor clave que representa un atributo y un valor clave que representa un valor numérico.

50 Además, el módulo 10 de detección configurado para seleccionar una parte de los pares de clave-valor a ser procesados como pares de clave-valor de punto caliente, está configurado específicamente para:
seleccionar aleatoriamente varios pares de clave-valor a ser procesados como pares de clave-valor de punto caliente. Además, el módulo 10 de detección configurado para seleccionar una parte de los pares de clave-valor a ser procesados como pares de clave-valor de punto caliente, está configurado específicamente para:
55 seleccionar aleatoriamente un primer número de pares de clave-valor a ser procesados como pares de clave-valor candidatos;
contar una frecuencia a la que se llama a cada uno de los pares de clave-valor entre los pares de clave-valor candidatos;
disponer los pares de clave-valor candidatos de acuerdo con las frecuencias; y
seleccionar un segundo número de pares de clave-valor que tienen frecuencias de llamada máximas a partir de los pares de clave-valor candidatos como pares de clave-valor de punto caliente;
60 en donde el primer número es mayor que el segundo número.

Además, el módulo 10 de detección configurado para seleccionar una parte de los pares de clave-valor de asignación como pares de clave-valor de punto caliente, está configurado además específicamente para:
65 establecer un conjunto de condiciones de categoría de servicio de pares de clave-valor candidatos antes del paso de seleccionar aleatoriamente un primer número de los pares de clave-valor a ser procesados como pares de clave-valor candidatos; y

seleccionar pares de clave-valor a ser procesados que cumplan con el conjunto de condiciones de categoría de servicio.

Además, el módulo 10 de detección configurado para seleccionar una parte de los pares de clave-valor a ser procesados como pares de clave-valor de punto caliente, está configurado específicamente para:

5 establecer un umbral de frecuencia de llamada de los pares de clave-valor de punto caliente; y
cuando la frecuencia de un par de clave-valor que se llama es mayor que el umbral de frecuencia de llamada, establecer el par de clave-valor como un par de clave-valor de punto caliente.

Además, el sistema incluye además un módulo 40 de optimización de reglas, configurado para:

10 optimizar la regla de detección utilizando un modelo de aprendizaje automático.

Además, el módulo 20 de asignación está configurado para:

asignar los pares de clave-valor que no son de punto caliente para obtener pares de clave-valor de resultado intermedio correspondientes a los pares de clave-valor que no son de punto caliente.

15 Además, un sistema 1 de procesamiento de datos incluye:

un módulo 20 de asignación configurado para asignar pares de clave-valor a ser procesados para obtener pares de clave-valor de resultado intermedio correspondientes a los pares de clave-valor a ser procesados;

un módulo 10 de detección configurado para seleccionar una parte de los pares de clave-valor de resultado intermedio como pares de clave-valor de punto caliente de acuerdo con una regla de detección; y

20 un módulo 30 de reducción configurado para reducir los pares de clave-valor de punto caliente para generar pares de clave-valor de resultado final para la llamada;

en donde el par de clave-valor incluye un valor clave que representa un atributo y un valor clave que representa el contenido del atributo.

25 En la realización de la presente solicitud, el sistema de procesamiento de datos procesa previamente los pares de clave-valor de punto caliente para facilitar la llamada por un sistema de servicio, mientras que los pares de clave-valor que no son de punto caliente se procesan solo cuando se llaman por el sistema de servicio, lo que reduce el volumen de datos que el sistema de procesamiento de datos debe procesar en tiempo real al proporcionar un servicio servidor para el sistema de servicio, mejora la eficiencia de ejecución del procesamiento de datos, reduce el tiempo del sistema de servicio esperando un resultado de procesamiento de datos y tiene un procesamiento de servicio fluido y una experiencia de usuario deseable.

30 de servicio esperando un resultado de procesamiento de datos y tiene un procesamiento de servicio fluido y una experiencia de usuario deseable.

Los expertos en la técnica deben comprender que las realizaciones de la presente invención pueden proporcionarse como un método, un sistema o un producto de programa informático. Por lo tanto, la presente invención puede implementarse en forma de una realización de hardware completa, una realización de software completa o una realización que combina software y hardware. Además, la presente invención puede ser un producto de programa informático implementado en uno o más medios de almacenamiento utilizables por computadora (que incluyen, entre otros, una memoria de disco magnético, un CD-ROM, una memoria óptica y similares) que incluye código de programa utilizable por computadora.

35 implementarse en forma de una realización de hardware completa, una realización de software completa o una realización que combina software y hardware. Además, la presente invención puede ser un producto de programa informático implementado en uno o más medios de almacenamiento utilizables por computadora (que incluyen, entre otros, una memoria de disco magnético, un CD-ROM, una memoria óptica y similares) que incluye código de programa utilizable por computadora.

40 La presente invención se describe con referencia a diagramas de flujo y/o diagramas de bloques de acuerdo con el método, dispositivo (sistema) y producto de programa informático de acuerdo con las realizaciones de la presente invención. Debe entenderse que se puede utilizar una instrucción de programa informático para implementar cada uno de los procesos y/o bloques en los diagramas de flujo y/o diagramas de bloques y combinaciones de procesos y/o bloques en los diagramas de flujo y/o diagramas de bloques. Estas instrucciones de programa informático pueden proporcionarse para una computadora de propósito general, una computadora de propósito especial, un procesador incorporado o un procesador de otro dispositivo de procesamiento numérico programable para generar una máquina, de modo que las instrucciones ejecutadas por una computadora o un procesador de otro dispositivo de procesamiento numérico programable generan un aparato para implementar una función específica en uno o más procesos en los diagramas de flujo y/o en uno o más bloques en los diagramas de bloques.

45 bloques en los diagramas de flujo y/o diagramas de bloques. Estas instrucciones de programa informático pueden proporcionarse para una computadora de propósito general, una computadora de propósito especial, un procesador incorporado o un procesador de otro dispositivo de procesamiento numérico programable para generar una máquina, de modo que las instrucciones ejecutadas por una computadora o un procesador de otro dispositivo de procesamiento numérico programable generan un aparato para implementar una función específica en uno o más procesos en los diagramas de flujo y/o en uno o más bloques en los diagramas de bloques.

50 Estas instrucciones de programa informático también pueden almacenarse en una memoria legible por computadora que puede instruir a la computadora u otro dispositivo de procesamiento numérico programable para que trabaje de una manera particular, de manera que las instrucciones almacenadas en la memoria legible por computadora generen un artículo de fabricación que incluya un aparato de instrucción. El aparato de instrucción implementa una función especificada en uno o más procesos en los diagramas de flujo y/o en uno o más bloques en los diagramas de bloques.

55 un artículo de fabricación que incluya un aparato de instrucción. El aparato de instrucción implementa una función especificada en uno o más procesos en los diagramas de flujo y/o en uno o más bloques en los diagramas de bloques.

60 Estas instrucciones del programa informático también pueden cargarse en una computadora u otro dispositivo de procesamiento numérico programable, de manera que se realicen una serie de pasos operativos en la computadora u otro dispositivo programable, generando así el procesamiento implementado por computadora. Por lo tanto, las instrucciones ejecutadas en la computadora u otro dispositivo programable proporcionan pasos para implementar una función especificada en uno o más procesos en los diagramas de flujo y/o en uno o más bloques en los diagramas de bloques.

65 En una configuración típica, un dispositivo informático incluye uno o más procesadores (CPU), una interfaz de entrada/salida, una interfaz de red y una memoria.

En una configuración típica, un dispositivo informático incluye uno o más procesadores (CPU), una interfaz de entrada/salida, una interfaz de red y una memoria.

La memoria puede incluir una memoria volátil, una memoria de acceso aleatorio (RAM) y/o una memoria no volátil o similar en un medio legible por computadora, por ejemplo, una memoria de solo lectura (ROM) o una RAM flash. La memoria es un ejemplo del medio legible por computadora.

5 El medio legible por computadora incluye medios no volátiles y volátiles, así como medios móviles y no móviles, y puede implementar el almacenamiento de información mediante cualquier método o tecnología. La información puede ser una instrucción legible por computadora, una estructura de datos y un módulo de un programa u otros datos. Un medio de almacenamiento de una computadora incluye, por ejemplo, pero no se limita a, una memoria de cambio de fase (PRAM), una memoria estática de acceso aleatorio (SRAM), una memoria dinámica de acceso aleatorio (DRAM), otros tipos de RAM, una ROM, una memoria de solo lectura programable y borrrable eléctricamente (EEPROM), una memoria flash u otras tecnologías de memoria, una memoria de solo lectura de disco compacto (CD-ROM), un disco versátil digital (DVD) u otros almacenamientos ópticos, una cinta de casete, un almacenamiento de cinta magnética/disco magnético u otros dispositivos de almacenamiento magnético, o cualquier otro medio sin transmisión, y puede utilizarse para almacenar información accesible al dispositivo informático. De acuerdo con la definición de este texto, el medio legible por computadora no incluye medios transitorios, tales como una señal de datos modulada y una portadora.

20 Cabe señalar además que, el término “incluye”, “comprende” u otras variaciones de los mismos están destinados a cubrir la inclusión no exclusiva, de modo que un proceso, método, producto o dispositivo que incluye una serie de elementos no solo incluye los elementos, sino que también incluye otros elementos que no están claramente enumerados, o incluye elementos inherentes del proceso, método, producto o dispositivo. Sin más limitaciones, un elemento definido por “incluye un/una...” no excluye que el proceso, método, producto o dispositivo que incluye el elemento tenga además otros elementos idénticos.

25 Los expertos en la técnica deben comprender que las realizaciones de la presente solicitud pueden proporcionarse como un método, un sistema o un producto de programa informático. Por lo tanto, la presente solicitud puede implementarse en la forma de una realización de hardware completa, una realización de software completa o una realización que combina software y hardware. Además, la presente solicitud puede emplear la forma de un producto de programa informático implementado en uno o más medios de almacenamiento utilizables por computadora (que incluyen, pero no se limitan a, una memoria de disco magnético, un CD-ROM, una memoria óptica y similares) que incluyen código de programa utilizable por computadora.

35 Las descripciones anteriores son meramente realizaciones de la presente solicitud, y no pretenden limitar la presente solicitud. Para los expertos en la técnica, la presente solicitud puede tener diversas modificaciones y variaciones.

REIVINDICACIONES

1. Un método de procesamiento de datos, el método que comprende:
 procesar previamente una pluralidad de pares de clave-valor de punto caliente, que comprende:
 5 seleccionar (S101) una primera pluralidad de pares de clave-valor de una pluralidad de pares de clave-valor a ser procesados, en donde la primera pluralidad se selecciona aleatoriamente como un conjunto candidato de pares de clave-valor para determinar la pluralidad de pares de clave-valor de punto caliente, en donde un par de clave-valor comprende un valor clave que representa un atributo y un valor clave que representa el contenido del atributo, en donde la primera pluralidad de pares de clave-valor corresponde a un primer servicio proporcionado por un sistema
 10 de servicio, en donde las llamadas incluyen uno o más pares de clave-valor de la pluralidad a ser procesada de pares de clave-valor que se reciben en un sistema (1) de procesamiento de datos desde el sistema de servicio;
 contar (S102) una frecuencia a la que el sistema de servicio llama a un par de clave-valor entre el conjunto de candidatos dentro de un período de tiempo;
 disponer (S103) la primera pluralidad de pares de clave-valor de acuerdo con la frecuencia;
 15 seleccionar (S104) una segunda pluralidad de pares de clave-valor del conjunto candidato de pares de clave-valor como la pluralidad de pares de clave-valor de punto caliente de acuerdo con una regla de detección, en donde la segunda pluralidad de pares de clave-valor es con frecuencias de llamada máximas, y en donde la primera pluralidad es mayor que la segunda pluralidad de pares de clave-valor;
 asignar (S200) la pluralidad de pares de clave-valor de punto caliente para obtener pares de clave-valor de resultado intermedio correspondientes a los pares de clave-valor de punto caliente; y
 20 reducir (S300) los pares de clave-valor de resultado intermedio de la pluralidad de pares de clave-valor de punto caliente para generar pares de clave-valor de resultado final para llamadas recibidas desde el sistema de servicio; el método que comprende, además:
 procesar una llamada desde el sistema de servicio en el sistema de procesamiento de datos, la llamada que incluye
 25 un par de clave-valor de la pluralidad a ser procesada de pares de clave-valor que no se selecciona como parte de la pluralidad de pares de clave-valor de punto caliente, que comprende:
 asignar el par de clave-valor para obtener un par de clave-valor de resultado intermedio correspondiente al par de clave-valor; y
 reducir el par de clave-valor de resultado intermedio para generar un par de clave-valor de resultado final para las
 30 llamadas desde el sistema de servicio.
2. El método de la reivindicación 1, en donde seleccionar (S104) la segunda pluralidad de pares de clave-valor como la pluralidad de pares de clave-valor de punto caliente de acuerdo con la regla de detección comprende, además:
 35 establecer un conjunto de condiciones de categoría de servicio de pares de clave-valor candidatos antes del paso de seleccionar aleatoriamente la primera pluralidad de pares de clave-valor de la pluralidad a ser procesada de pares de clave-valor como pares de clave-valor candidatos; y
 seleccionar pares de clave-valor a ser procesados de la pluralidad a ser procesada de pares de clave-valor que cumplen el conjunto de condiciones de categoría de servicio.
3. El método de la reivindicación 1, en donde seleccionar (S104) la segunda pluralidad de pares de clave-valor como la pluralidad de pares de clave-valor de punto caliente de acuerdo con la regla de detección comprende:
 40 establecer un umbral de frecuencia de llamada de los pares de clave-valor de punto caliente; y
 cuando la frecuencia de un par de clave-valor que se llama es mayor que el umbral de frecuencia de llamada, establecer el par de clave-valor como un par de clave-valor de punto caliente.
4. El método de una cualquiera de las reivindicaciones 1 a 3, en donde el método comprende, además optimizar la regla de detección utilizando un modelo de aprendizaje automático, que opcionalmente comprende:
 45 contar una condición de distribución de las frecuencias de los pares de clave-valor que se llaman frente a un solo atributo utilizando un algoritmo de agrupamiento en el modelo de aprendizaje automático;
 50 de acuerdo con la condición de distribución de frecuencias de los pares de clave-valor que se llaman frente a un solo atributo, seleccionar un intervalo de valores clave del contenido del atributo en el que las frecuencias de los pares de clave-valor que se llaman no son inferiores a un umbral de frecuencia preestablecido; y
 establecer el intervalo de valores clave del contenido del atributo como una condición de la regla de detección.
5. El método de la reivindicación 4, en donde el método comprende, además:
 55 cuando un par de clave-valor de un atributo y un par de clave-valor de otro atributo se llaman por sistemas de servicio que tienen un mismo código de servicio, establecer una unión de conjunto de intervalos los de valores clave del contenido del atributo de los pares de clave-valor de los dos atributos como una condición de la regla de detección.
6. Un sistema (1) de procesamiento de datos, que comprende:
 60 un módulo (10) de detección configurado, en un modo de procesamiento previo, para
 seleccionar (S101) una primera pluralidad de pares de clave-valor de una pluralidad a ser procesada de pares de clave-valor, en donde la primera pluralidad se selecciona aleatoriamente como un conjunto candidato de pares de clave-valor para determinar una pluralidad de pares de clave-valor de punto caliente, en donde la primera pluralidad
 65 de pares de clave-valor corresponde a un primer servicio proporcionado por un sistema de servicio, en donde las

- llamadas que incluyen uno o más pares de clave-valor de la pluralidad a ser procesada de pares de clave-valor que se reciben en un sistema de procesamiento de datos desde el sistema de servicio;
 contar (S102) la frecuencia a la que se llama a un par de clave-valor entre el conjunto de candidatos dentro de un período de tiempo mediante el sistema de servicio (S102);
 5 disponer (S103) la primera pluralidad de pares de clave-valor de acuerdo con la frecuencia (S103);
 seleccionar (S104) una segunda pluralidad de pares de clave-valor del conjunto candidato de pares de clave-valor como la pluralidad de pares de clave-valor de punto caliente de acuerdo con una regla de detección, en donde la segunda pluralidad de pares de clave-valor es con frecuencias de llamada máximas;
 10 un módulo (20) de asignación configurado, en el modo de procesamiento previo, para asignar (S200) la pluralidad de pares de clave-valor de punto caliente para obtener pares de clave-valor de resultado intermedio correspondientes a los pares de clave-valor de punto caliente; y
 un módulo (30) de reducción configurado, en el modo de procesamiento previo, para reducir (S300) los pares de clave-valor de resultado intermedio de la pluralidad de pares de clave-valor de punto caliente para generar pares de clave-valor de resultado final para llamadas recibidas desde el sistema de servicio;
 15 en donde el par de clave-valor comprende un valor clave que representa un atributo y un valor clave que representa el contenido del atributo;
 en donde el módulo (20) de asignación está configurado, en un modo de procesamiento, para procesar una llamada desde el sistema de servicio, la llamada que incluye un par de clave-valor de la pluralidad a ser procesada de pares de clave-valor que no se selecciona como siendo parte de la pluralidad de pares de clave-valor de punto caliente y
 20 asignar el par de clave-valor que no es de punto caliente para obtener un par de clave-valor de resultado intermedio correspondiente al par de clave-valor; y
 en donde el módulo (30) de reducción está configurado, en el modo de procesamiento, para reducir el par de clave-valor de resultado intermedio para generar un par de clave-valor de resultado final para las llamadas desde el sistema de servicio.
- 25 7. El sistema (1) de la reivindicación 6, en donde el módulo (10) de detección configurado para seleccionar la pluralidad de pares de clave-valor como la pluralidad de pares de clave-valor de punto caliente, está configurado además para:
 establecer un conjunto de condiciones de categoría de servicio de pares de clave-valor candidatos antes del paso de seleccionar aleatoriamente la primera pluralidad de pares de clave-valor de la pluralidad a ser procesada de pares de
 30 clave-valor como pares de clave-valor candidatos; y
 seleccionar los pares de clave-valor a ser procesados de la pluralidad a ser procesada de pares de clave-valor que cumplan el conjunto de condiciones de categoría de servicio.
- 35 8. El sistema (1) de la reivindicación 6, en donde el módulo (10) de detección configurado para seleccionar la segunda pluralidad de pares de clave-valor como la pluralidad de pares de clave-valor de punto caliente, está configurado para:
 establecer un umbral de frecuencia de llamada de los pares de clave-valor de punto caliente; y
 cuando la frecuencia de un par de clave-valor que se llama es mayor que el umbral de frecuencia de llamada, establecer el par de clave-valor como un par de clave-valor de punto caliente.
- 40 9. El sistema (1) de una cualquiera de las reivindicaciones 6 a 8, en donde el sistema comprende, además un módulo (40) de optimización de reglas configurado para:
 optimizar la regla de detección utilizando un modelo de aprendizaje automático.
- 45 10.. El sistema (1) de la reivindicación 9, en donde el módulo (40) de optimización de reglas está configurado para:
 contar una condición de distribución de las frecuencias de los pares de clave-valor que se llaman frente a un solo atributo utilizando un algoritmo de agrupamiento en el modelo de aprendizaje automático;
 de acuerdo con la condición de distribución de las frecuencias de los pares de clave-valor que se llaman frente a un solo atributo, seleccionar un intervalo de valores clave del contenido del atributo en el que las frecuencias de los pares
 50 de clave-valor que se llaman no sean inferiores a un umbral de frecuencia preestablecido; y
 establecer el intervalo de valores clave del contenido del atributo como una condición de la regla de detección.
- 55 11. El sistema (1) de la reivindicación 10, en donde el módulo (40) de optimización de reglas está configurado además para:
 cuando un par de clave-valor de un atributo y un par de clave-valor de otro atributo se llaman mediante los sistemas de servicio que tienen un mismo código de servicio, establecer una unión de conjuntos de intervalos de valores clave del contenido del atributo de los pares de clave-valor de los dos atributos como una condición de la regla de detección.

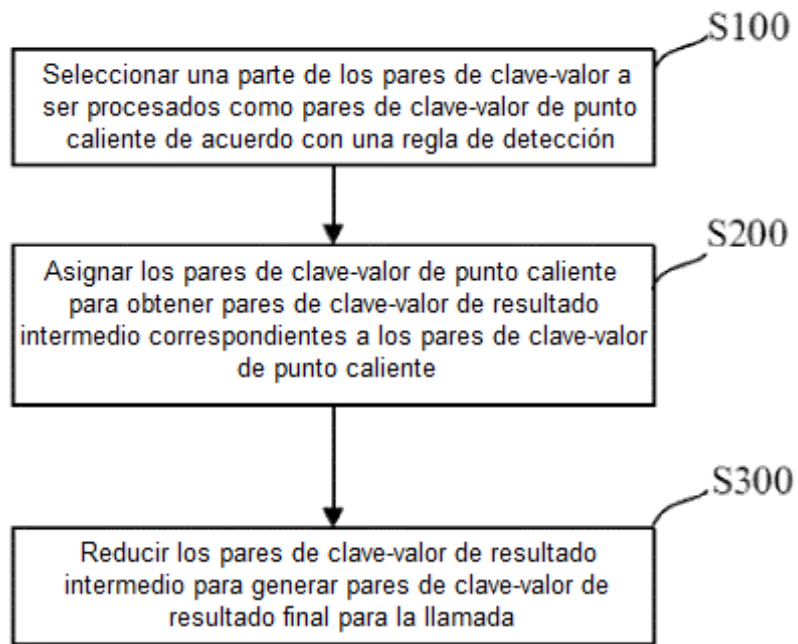


FIG. 1

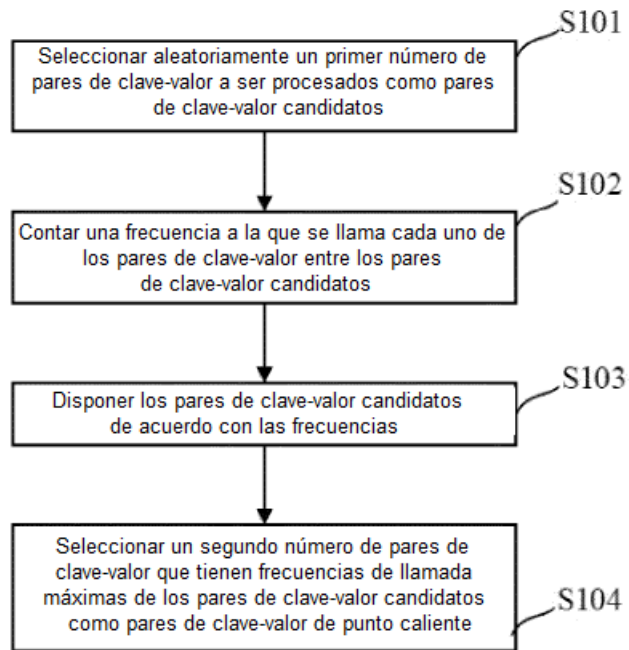


FIG. 2

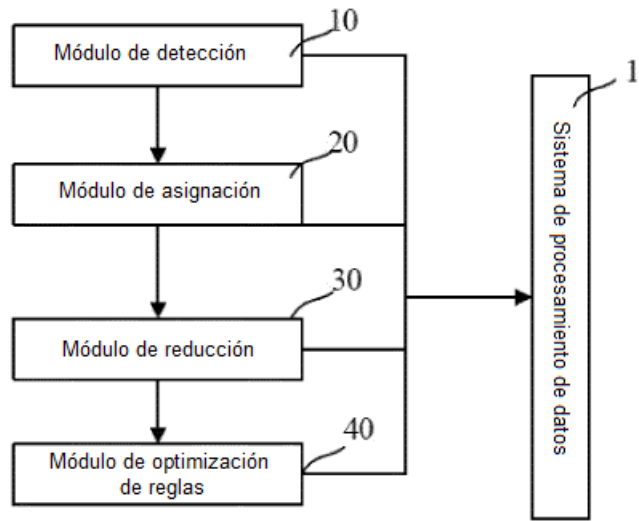


FIG. 3