

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 743 803**

51 Int. Cl.:

**G06F 17/30** (2006.01)

**G10L 15/22** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **04.04.2014** E 14163510 (2)

97 Fecha y número de publicación de la concesión europea: **29.05.2019** EP 2927818

54 Título: **Procedimiento para el procesamiento automático de un número de ficheros de registro de un sistema de automatización**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**20.02.2020**

73 Titular/es:  
**SIEMENS AKTIENGESELLSCHAFT (100.0%)**  
**Werner-von-Siemens-Straße 1**  
**80333 München, DE**

72 Inventor/es:  
**STORL, ALEXANDER;**  
**PAUER, STEFFEN y**  
**ROSHCHIN, MIKHAIL**

74 Agente/Representante:  
**CARVAJAL Y URQUIJO, Isabel**

ES 2 743 803 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**

Procedimiento para el procesamiento automático de un número de ficheros de registro de un sistema de automatización

5 La invención se relaciona con un procedimiento para el procesamiento automático de un gran número de ficheros de registro de un sistema de automatización de una instalación técnica, como una planta de energía.

Las instalaciones a gran escala, como las plantas de turbinas de vapor y de gas, las plantas químicas, las acerías, las laminadoras, son sistemas técnicos complejos, que tienen un gran número de componentes (por ejemplo, varios cientos), como sensores, actuadores, accionamientos, etc., de los que, en su caso, cada uno recopila y proporciona múltiples lecturas por segundo para su posterior procesamiento y transmisión.

10 Para la monitorización y control de componentes, como turbinas, evaporadores, generadores, sensores, accionamientos, actuadores de la instalación técnica, se procesan, analizan e interpretan los datos obtenidos por los componentes. Así es posible identificar lo antes posible desviaciones de un estado normal predeterminado y, si fuera necesario, evitar una avería y/o un fallo del componente, como la turbina.

15 En este contexto, la cantidad de datos/señales de medición a evaluar y la complejidad de las posibles dependencias entre los datos/señales de medición son habitualmente demasiado grandes, como para que un operador pueda realizar un análisis efectivo de los datos/señales de medición.

20 La instalación técnica comprende además para el control y monitorización de la instalación un sistema de automatización, que generalmente se subdivide en un nivel de operación y de monitorización, un nivel de automatización/ proceso y un nivel de campo, que están conectados entre sí a través de una unidad de comunicaciones, particularmente uno o varios sistemas de buses de datos.

25 Los datos, como señales de medición, señales de estado de los componentes individuales, como sensores, actuadores, accionamientos, turbinas, unidades de automatización, etc., se resumen habitualmente en ficheros de registro, que describen uno o varios eventos ocurridos de uno o varios componentes y/o grupos de componentes, y se transmiten a los componentes afectados de la instalación técnica. La patente WO00/39674A1 revela un administrador de eventos (Event Manager) y un correlacionador de eventos (Event Korrelator) y la determinación de correlaciones de eventos.

30 En la práctica del control y/o monitorización automático/a existe el problema del número creciente de ficheros de registro generados al aparecer un evento en la instalación técnica, por ejemplo, cuando ocurre un accidente. Las causas típicas de esto son, por ejemplo, sensores defectuosos y/o piezas defectuosas de la instalación, que proporcionan valores medidos adicionales y, por consiguiente, ficheros de registro adicionales. Esto dificulta al personal operativo la visibilidad y la detección de los ficheros de registro generados. Además, es difícil para el personal operativo evaluar correctamente el evento ocurrido.

Por otra parte, se genera un gran número de ficheros de registro diferentes, cuyos contenidos no son compatibles ni comparables.

35 El objeto de la invención consiste en especificar un sencillo procedimiento para el procesamiento automático de un gran número de ficheros de registro de un sistema de automatización de una instalación técnica.

Este objeto se resuelve según las características de las reivindicaciones independientes. Las formas de ejecución preferidas pueden extraerse particularmente de las reivindicaciones dependientes.

40 Por ejemplo, para los caracteres especiales contenidos en un juego de datos o caracteres no numéricos o no alfabéticos, se insertan automáticamente espacios, donde para dos o varios espacios consecutivos resultantes de esto, se eliminará del juego de datos un número de espacios consecutivos superior a uno. De este modo se extraerán aquellos caracteres en la parte del mensaje (por ejemplo, parte del texto, parte del mensaje, parte del estado) del fichero de registro, que no describan el evento con más detalle, como los caracteres especiales, por ejemplo, ":", "+", "\$", "#", etc., de esta parte del mensaje y, por consiguiente, la parte del mensaje de la parte del sistema (por ejemplo, hora del sistema, versión, revisión) del fichero de registro. Por consiguiente, el análisis del fichero de registro se limita a la parte del mensaje relevante para el contenido y, por lo tanto, se simplifica y acelera significativamente.

45 Además, la concatenación de los contenidos del respectivo juego de datos se resuelve de forma que los contenidos se separen nuevamente. Particularmente se subdividen los contenidos además en palabras. Los contenidos alfanuméricos separados resultantes de la resolución de la concatenación, particularmente las palabras con menos

de cuatro caracteres, se eliminan para la reducción de los caracteres/datos a seguir procesando y, por consiguiente, para la compresión de datos. Las palabras con menos de cuatro caracteres son irrelevantes en el análisis, pues sólo representan un "ruido estadístico" para los cálculos de similitud posteriores y, por consiguiente, no son relevantes.

5 Un posible modo de operación de la codificación prevé que los juegos de datos, particularmente sus contenidos alfanuméricos separados, se analicen y estandaricen de tal forma que a cada contenido alfanumérico separado se le asigne una palabra de referencia y a cada palabra de referencia, un código numérico. Alternativamente al código numérico puede emplearse un código alfabético y/o un código alfanumérico. El código numérico tiene la ventaja de una clasificación y agrupación sencilla y rápida de los juegos de datos de los ficheros de registro.

10 Para estandarizar y comprimir aún más los juegos de datos de diferentes ficheros de registro y el análisis sencillo y rápido de estos juegos de datos, el archivo resumen con los contenidos alfanuméricos separados de todos los ficheros de registro se convierte en base a las palabras de referencia determinadas y/o del/de los código/s numérico/s en una matriz de códigos bidimensional, particularmente alfanumérica.

15 En un modo de operación especialmente sencillo, los códigos numéricos de la matriz de códigos bidimensional se ordenan cronológicamente y/o en función de la respectiva longitud de caracteres. Además, códigos numéricos similares de la matriz de códigos pueden agruparse por medio de una operación de similitud, particularmente mediante la llamada distancia de Levenshtein, particularmente asignarse a al menos un grupo. Además, cada grupo se describe o representa mediante un código de grupo asociado.

20 Así puede, para una determinada distancia de Levenshtein de cero para un nuevo código numérico a considerar respecto a uno de los grupos, asignarse este código numérico a considerar a este grupo, pues el código numérico a considerar es idéntico al grupo, particularmente al código de grupo que representa a este grupo.

Por el contrario, para una distancia de Levenshtein de uno para un código numérico a considerar respecto a uno de los grupos y para una diferencia del código numérico a considerar sólo en una posición de dígito a este grupo, este código numérico a considerar se asigna a este grupo.

25 Por otra parte, puede compararse cada nuevo código numérico a considerar con códigos numéricos no agrupados de la misma longitud o con una longitud mayor de como máximo un carácter.

Para el caso de que la distancia de Levenshtein de un nuevo código numérico a considerar respecto a otro código numérico no-agrupado sea igual a uno, se forma un nuevo grupo con un nuevo código de grupo a partir de ambos códigos numéricos de tal forma que

30 - para una longitud diferente de ambos códigos numéricos a considerar, se asigne al nuevo grupo el código numérico más corto como nuevo código de grupo o

- para la misma longitud de ambos códigos numéricos con al menos un carácter diferente del nuevo grupo de códigos numéricos, sin ese carácter diferente se asigna como nuevo código de grupo.

35 Adicional o alternativamente, cuando se identifique un nuevo código numérico a considerar, que no tenga una distancia de Levenshtein de igual a uno respecto a otro código numérico no-agrupado, este código numérico a considerar puede asignarse a otro nuevo grupo. Además, a este otro nuevo grupo se le asigna este nuevo código numérico a considerar como nuevo código de grupo. Los futuros nuevos códigos numéricos a considerar, que, en comparación con este nuevo código de grupo tienen una distancia de Levenshtein de uno, se asignan a este nuevo grupo.

40 Por consiguiente, la matriz de códigos bidimensional determinada y los grupos determinados con códigos de grupo para agrupar códigos numéricos similares y/o idénticos, que describan un evento, pueden extenderse y complementarse continuamente agregando nuevos códigos numéricos de los nuevos juegos de datos de los nuevos ficheros de registro en la forma de un sistema experto.

45 Por otra parte, el procedimiento permite que, al iniciar el procedimiento, sólo se emplee un número predeterminado de grupos y no se generen nuevos grupos. De este modo se posibilita primero una agrupación robusta de los nuevos códigos numéricos a considerar. En la operación continua puede permitirse entonces la generación de nuevos grupos.

El procedimiento descrito se caracteriza porque las partes del mensaje, inicialmente no compatibles y no comparables de los diversos ficheros de registro mediante la concatenación de todos los componentes en el estado original y posterior compresión de estos mediante reducción de caracteres y codificación de esta cadena reducida,

así como agrupación de cadenas de caracteres codificadas similares o idénticas, se clasifican y analizan mucho más rápido respecto a los procedimientos de análisis convencionales.

Preferentemente, el procedimiento se usa en un sistema de automatización de una instalación técnica, como una planta de energía.

- 5 En un sencillo modo de operación, el procedimiento conforme a la invención se implementa en un componente de un sistema de automatización de una instalación técnica, donde el sistema de automatización comprende al menos un nivel de operación y monitorización, un nivel de automatización/proceso y un nivel de campo, interconectados a través de un sistema de comunicaciones.

- 10 Las propiedades, características y ventajas descritas anteriormente de esta invención, así como la manera en que se logran, se harán más claras y se entenderán más claramente en relación con la siguiente descripción de ejemplos de ejecución, que se explican con más detalle en base a los dibujos. Además, muestran:

FIG 1 esquemáticamente un ejemplo de ejecución para un sistema de automatización para el control y monitorización de una instalación técnica, y

- 15 FIG 2 esquemáticamente un ejemplo de ejecución para un procedimiento para el procesamiento de ficheros de registro, generados en el sistema de automatización.

Las piezas mutuamente correspondientes están provistas en todas las Figuras de los mismos símbolos de referencia.

La Figura 1 muestra esquemáticamente esquemático en diagrama de bloques un ejemplo de ejecución para un sistema de automatización 1 de una instalación técnica, como una planta de energía o una planta química.

- 20 El sistema de automatización 1 comprende, en un nivel de proceso AS (también denominado nivel de automatización), un número de aparatos de automatización 2, interconectados por un bus de datos 3 común y que están conectados y comunican con una unidad deflectora 4 en un nivel de operación y de monitorización BB. Los aparatos de automatización 2 están conectado, por otra parte, a través de un bus de campo 5 con aparatos de campo 6 en un nivel de campo FE y se comunican entre sí.

- 25 El bus de datos 3 puede ser, por ejemplo, un bus de Ethernet u otra unidad de transmisión de datos adecuada. El bus de campo 5 puede ser, por ejemplo, un bus de campo basado en Ethernet 5 o estar diseñado como otra apropiada unidad de transmisión de datos inalámbrica o cableada. El bus de datos 3 y/o el bus de campo 5 pueden formar además una red informática.

- 30 Los aparatos de automatización 2 pueden estar compuestos tanto por procesadores libremente programables como también programables por memoria, particularmente por una unidad de procesador y controlan, regulan y/o monitorizan en cada caso varios grupos de componentes 7 o subsistemas de los componentes individuales de la instalación técnica. Particularmente, el control, regulación y/o monitorización de la instalación técnica comprende también un análisis y otro procesamiento de los datos, como señales de medición, señales de control, señales de entrada y salida, señales intermedias, datos almacenados, datos de procesamiento.

- 35 Las unidades deflectoras 4 pueden ser unidades de procesamiento de datos, como un ordenador personal u otra unidad de control y monitorización adecuada, por medio de la cual se monitorice y controle la instalación técnica.

- 40 Los aparatos de campo 6 pueden ser, por ejemplo, controladores lógicos compactos libremente programables o programables por memoria, particularmente una unidad de procesador y/o al menos un sistema de conexión parcialmente cableado o lógico, que controlen, regulen y/o monitoricen componentes individuales 8 de los componentes del sistema, tales como sondas de medición (sensores) y / o accionadores (actuadores) , y que estén conectados a través del bus de campo 5 con el fin de comunicarse con el aparato de automatización 2 (también denominado aparato de control).

- 45 Durante la operación de una instalación técnica, como una planta de energía, a través del bus de datos 3 y el bus de campo 5 se mueven grandes cantidades de datos en forma de ficheros de registro P1 a Pn, que incluyen comandos de control, mensajes de estado y/o mensajes de fallo, activados en la unidad deflectora 4 mediante intervenciones de control y/o pueden hacerse visibles sobre pantallas u otros aparatos de visualización de la unidad deflectora 4. Además, particularmente durante la puesta en marcha y al iniciar/detener la instalación, se precisa un gran número de intervenciones de control, que tienen como consecuencia un número correspondientemente grande de retroalimentaciones a considerar. Esto conlleva una carga considerable sobre el personal operativo.

La Figura 2 muestra esquemáticamente un ejemplo de ejecución para un procedimiento para el procesamiento, particularmente preprocesamiento, de ficheros de registro P1 a Pn, generados en el sistema de automatización 1 y transmitidos a través del bus de datos 3 y/o del bus de campo 5 e intercambiados entre las unidades del sistema de automatización 1, particularmente entre las unidades deflectoras 4, los aparatos de automatización 2 y los aparatos de campo 6.

Los ficheros de registro P1 a Pn incluyen en cada caso al menos una fecha de detección y un texto, particularmente un mensaje, un aviso, un texto de estado, que describan al menos un evento. Además, los ficheros de registro P1 a Pn pueden incluir como texto una hora del sistema, un número de versión / revisión y otra información. Además, por ejemplo, varios ficheros de registro P1 a Pn pueden describir uno o varios eventos. Como ficheros de registro P1 a Pn se producen y generan, por ejemplo, en un sistema de automatización 1 de una instalación técnica varios archivos \*log-, \*xml-, \*txt-, \*nfo- y/o \*dmp-.

En al menos uno de los componentes del sistema de automatización 1 se implementa particularmente un producto de programa informático, que se pueda cargar directamente en una memoria de un ordenador digital, como una unidad deflectora 4, comprendiendo partes de código de programa, apropiadas para llevar a cabo pasos del procedimiento descrito a continuación. Alternativamente, el producto del programa informático puede cargarse también en un aparato de automatización 2.

Alternativamente puede preverse un medio de almacenamiento legible por computadora, por ejemplo, cualquier memoria, que comprenda instrucciones ejecutables por un ordenador, como una unidad deflectora 4 (por ejemplo, en forma de códigos de programa), apropiadas para que la unidad deflectora 4 realice los pasos del procedimiento descrito a continuación.

El procedimiento implementado como un compresor de registros 9 para el procesamiento automático, particularmente el preprocesamiento, del gran número de ficheros de registro P1 a Pn generados, particularmente heterogéneos o incompatibles, comprende al menos los siguientes pasos:

- Determinación de al menos una parte del mensaje N1 a Nn de cada juego de datos D1 a Dn de un fichero de registro P1 a Pn mediante:
- Concatenación de todos los contenidos del respectivo juego de datos D1 a Dn del fichero de registro P1 a Pn,
- Reunión de los juegos de datos D1 a Dn de todos los ficheros de registro P1 a Pn en un archivo resumen SD,
- Clasificación cronológica de los juegos de datos D1 a Dn de todos los ficheros de registro P1 a Pn en el archivo resumen SD,
- Compresión de cada juego de datos D1 a Dn en el archivo resumen SD,
- Codificación de los juegos de datos comprimidos kD1 a kDn en el archivo resumen SD mediante al menos uno o varios código(s) numérico(s) Kx,
- Ordenación y/o asignación, particularmente agrupación, de los juegos de datos comprimidos kD1 a kDn en base al/a los código(s) numérico(s) Kx determinado(s) a al menos un grupo Gy, que tenga al menos un código de grupo asociado KGy, y
- Emisión de la parte alfanumérica del mensaje N1 a Nn del fichero de registro P1 a Pn mediante decodificación del código de grupo KGy del grupo Gy asignado y emisión del código de grupo decodificado KGy.

La parte del mensaje N1 a Nn de cada juego de datos D1 a Dn de los ficheros de registro P1 a Pn incluye, por ejemplo, mensajes de texto, como notificaciones, estados, textos, información, advertencias, etc. El respectivo juego de datos D1 a Dn puede tener, por ejemplo, la forma de una tabla o una base de datos o tener otra forma adecuada con particiones de campo.

La parte del mensaje N1 a Nn del respectivo fichero de registro P1 a Pn puede estar contenida, por ejemplo, en uno o varios campos del juego de datos D1 a Dn asociado. Otros campos del juego de datos D1 a Dn contienen información del registro y/o del sistema, como, por ejemplo, la hora del sistema, el número de versión, el número de revisión, etc.

## ES 2 743 803 T3

El procedimiento conforme a la invención prevé que los contenidos de los campos del respectivo juego de datos D1 a Dn de cada fichero de registro P1 a Pn se concatenen juntos en un primer paso y a continuación todos los juegos de datos D1 a Dn concatenados de todos los ficheros de registro P1 a Pn se resuman en un archivo resumen SD, particularmente una única tabla o base de datos. Por ejemplo, dos ficheros de registro P1 y P2 incluyen los siguiente contenidos o entradas:

Fichero de registro P1:

28.03.2014 ABC

21.03.2014 ZDF

Fichero de registro P2:

27.03.2014 XYZ

01.01.2014 SAP

Después de concatenar y resumir los contenidos en el S archivo resumen SD, éste comprende los contenidos como sigue:

28.03.2014 ABC

27.03.2014 XYZ

21.03.2014 ZDF

01.01.2014 SAP.

Dentro de este archivo de suma SD particularmente en forma de tabla, los juegos de datos D1 a Dn se ordenan con los contenidos concatenados en función de una fecha respectivamente asociada, en particular cronológicamente, por ejemplo, en orden ascendente o descendente.

La fecha asociada puede ser particularmente una fecha de detección para el mensaje subyacente, que describa particularmente un evento como "temperatura ambiente en el rango permitido", "arranque inicial activado", "bomba fallida", "turbina iniciada", etc., en el anexo técnico. Alternativa o adicionalmente, los juegos de datos D1 a Dn pueden clasificarse en base a parámetros de registro alternativos o adicionales, tales como lugar de detección, tipo de detección y/o período de detección.

Durante la posterior compresión de los juegos de datos D1 a Dn con los contenidos concatenados, estos se procesan de tal manera que al menos la longitud de los datos y/o el contenido de los datos del respectivo juego de datos D1 a Dn se reduzca(n).

Por ejemplo, para los caracteres especiales contenidos en un en un juego de datos D1 a Dn, como por ejemplo "\$", "#", "@" etc., o caracteres no numéricos o no alfabéticos, "-", "/", etc. se utilizan automáticamente espacios. A continuación, en la cadena de contenido del respectivo juego de datos D1 a Dn, si hubiera dos o más espacios consecutivos, estos se reducen automáticamente a un espacio eliminando el número de espacios sucesivos superior a un solo espacio. De este modo, aquellos caracteres del juego de datos D1 a Dn, particularmente de la parte del mensaje N1 a Nn (por ejemplo, parte del texto, parte del mensaje, parte del estado) se eliminan del fichero de registro P1 a Pn, que no describan el evento en detalle. Por consiguiente, la parte del mensaje N1 a Nn se extrae de manera simple de la parte del sistema (por ejemplo, hora del sistema, versión, revisión) del fichero de registro P1 a Pn, de forma que el procesamiento y análisis ulteriores del fichero de registro se limite a la parte del mensaje relevante al contenido N1 a Nn y, por lo tanto, se simplifica y acelera significativamente.

En otro paso se resuelve entonces la concatenación de los contenidos de la parte del mensaje N1 a Nn, por lo cual los contenidos, particularmente las palabras/frases, se separan. Además, los contenidos se subdividen, por ejemplo, en palabras/frases.

Para la compresión adicional de datos de los juegos de datos D1 a Dn, se eliminan, por ejemplo, aquellas palabras/frases separadas del juego de datos D1 a Dn, que tengan menos de cuatro, particularmente menos de tres caracteres. Por ejemplo, se elimina la palabra "es" o "sí". Esto posibilita comparar entre sí de manera sencilla y efectiva dos o más entradas/contenidos diferentes sintácticamente (sin significado semántico) y automáticamente.

Además, se prevé que los juegos de datos D1 a Dn, particularmente sus contenidos alfanuméricos separados, se analicen y estandaricen como las palabras/frases separadas sobrantes con particularmente más de tres caracteres, de tal forma que a cada contenido alfanumérico separado se le asigne una palabra de referencia Rx y a cada palabra de referencia Rx, un código numérico Kx. De este modo se reducen significativamente las necesidades de almacenamiento para archivar los juegos de datos D1 a Dn y su análisis se acelera claramente.

5

Por ejemplo, las palabras y/o frases de los contenidos de mensaje:

- "temperatura ambiente dentro del rango permitido",

- "temperatura de la turbina en el rango verde",

- "primer inicio activado",

10 - "bomba fallida",

- " turbina iniciada"

se convierten en las siguientes palabras de referencia Rx

- "la temperatura ambiente está permitida",

- "la temperatura de la turbina está permitida",

15 - "primer inicio activado",

- "fallo de la bomba",

- "inicio de la turbina"

y / o en el siguiente código numérico Kx:

- "123",

20 - "423"

- "56",

- "7",

- "8",

25 Además, se codifican palabras individuales, por ejemplo, por medio de un código numérico de un solo dígito Kx. Las frases con más de una palabra se codifican, por ejemplo, por medio de un código numérico Kx, que presenta un número de dígitos correspondiente al número de palabras. Las palabras y/o frases idénticas se codifican con la misma palabra de referencia Rx y el mismo código numérico Kx. Las palabras y/o frases de diferentes juegos de datos D1 a Dn con caracteres y/o palabras parcialmente coincidentes se codifican con códigos numéricos Kx coincidentes al menos en esas posiciones de dígito.

30 Alternativamente al código numérico Kx puede usarse un código alfabético y/o un código alfanumérico. La codificación de los juegos de datos D1 a Dn por medio del código numérico Kx tiene la ventaja de una sencilla y rápida clasificación y agrupación de los juegos de datos D1 a Dn de los ficheros de registro P1 a Pn.

35 Para la estandarización y compresión adicional de los juegos de datos D1 a Dn de los diferentes ficheros de registro P1 a Pn y el sencillo y rápido análisis de estos juegos de datos D1 a Dn, el archivo resumen en forma de tabla SD con los contenidos alfanuméricos separados de todos los ficheros de registro P1 a Pn se convierte, en base a las palabras de referencia Rx determinadas y/o al código numérico Kx, en una matriz bidimensional, particularmente de códigos alfanuméricos KM.

En la matriz bidimensional de códigos KM se ordenan entonces los códigos numéricos Kx cronológicamente, particularmente de manera cronológicamente ascendente o descendente, y/o en función de la respectiva longitud de

caracteres. Por ejemplo, los códigos numéricos Kx "5632", "543", "64221", "123" se ordenan de la siguiente manera: "123", "543", "5632", "64221".

5 Además, en un paso adicional pueden agruparse códigos numéricos similares Kx de la matriz de códigos KM por medio de una operación de similitud, particularmente mediante la llamada distancia de Levenshtein, particularmente asignándose a al menos un grupo Gy (= clúster).

Además, cada grupo Gy se describe o representa a través de un código de grupo asociado KGy. El código de grupo asociado KGy puede generarse además a partir de al menos el código numérico Kx de un primer juego de datos D1 a Dn, asignado a este grupo Gy. Alternativamente puede especificarse el respectivo código de grupo KGy de uno o de varios grupos/s Gy.

10 Por ejemplo, para una determinada distancia de Levenshtein de cero para un nuevo código numérico a considerar Kx respecto a uno de los grupos Gy, puede asignársele este código numérico a considerar Kx a este grupo Gy, pues el código numérico a considerar Kx es idéntico al grupo Gy, particularmente a este código de grupo KGy que representa al grupo Gy.

15 Por el contrario, para una distancia de Levenshtein de uno para un código numérico a considerar Kx respecto a uno de los grupos Gy y para una diferencia del código numérico a considerar Kx de sólo una posición de un dígito respecto a este grupo Gy, este código numérico a considerar Kx se asigna a este grupo Gy, por ejemplo, en un grupo Gy con un código de grupo asociado KGy de "12", a este grupo Gy se le asignan los códigos numéricos Kx con las siguientes posiciones "123", "124", "12" y/o "13".

20 Para, por ejemplo, tanto en un posterior análisis como también en la agrupación de los códigos numéricos Kx, mejorar y acelerar las funciones de búsqueda, se inicia la comparación del nuevo código numérico a considerar Kx con los grupos Gy ya generados en el grupo Gy creado en último lugar o en el grupo Gy, al que se le ha asignado en último lugar un código numérico previo Kx.

25 En particular, el inicio de la agrupación o asignación de un nuevo código numérico Kx a considerar en el grupo Gy, al que se le ha asignado un código numérico previo Kx, acelera el algoritmo de agrupación, pues habitualmente los códigos numéricos Kx adyacentes, particularmente temporalmente posteriores, pero también temporalmente previos, que anteriormente se ordenaron temporalmente, se asignan a uno y el mismo evento y, por consiguiente, pueden asignarse a uno y el mismo grupo Gy.

30 Por otra parte, cada recién generado y, por consiguiente, nuevo código numérico a considerar Kx puede con compararse códigos numéricos no agrupados Kx de la misma longitud o con una mayor longitud de como máximo un carácter.

Para el caso de que la distancia de Levenshtein de un nuevo código numérico a considerar Kx respecto a otro código numérico no-agrupado Kx sea igual a uno, se forma un nuevo grupo Gy+1 con un nuevo código de grupo KGy+1 de ambos códigos numéricos Kx de tal forma, que

35 - para una longitud diferente de ambos códigos numéricos a considerar Kx, al nuevo grupo Gy+1 se le asigna el código numérico más corto Kx como nuevo código de grupo KGy+1 (por ejemplo, Kx = "123" y "12" conlleva un nuevo grupo Gy con nuevo código de grupo KGy+1 de "12") o

- para la misma longitud de ambos códigos numéricos Kx con al menos un carácter diferente del nuevo grupo Gy+1, se asigna el código numérico Kx sin el carácter diferente como nuevo código de grupo KGy+1 (por ejemplo, Kx = "133" y "134" conlleva un nuevo grupo Gy con el nuevo código de grupo KGy+1 de "13").

40 Adicional o alternativamente puede, cuando se identifique un nuevo código numérico Kx a considerar, que no presente una distancia de Levenshtein de igual a uno respecto a otro código numérico no-agrupado Kx, este código numérico a considerar Kx se asigna a otro nuevo grupo Gy+1. Además, a este otro nuevo grupo Gy+1 se le asigna este nuevo código numérico a considerar Kx como código de grupo KGy+1. Los nuevos códigos numéricos futuros a considerar Kx, que, en comparación con este nuevo código de grupo KGy+1, presentan una distancia de Levenshtein de uno, se asignan a este nuevo grupo Gy+1.

45

La matriz de códigos bidimensional KM representa, por consiguiente, un sistema dinámico de expertos, en que los grupos determinados Gy, Gy+1 con códigos de grupo KGy, KGy+1 para agrupar códigos numéricos Kx similares y/o idénticos, que describen un evento, se extienden y complementan continuamente mediante adición de nuevos códigos numéricos Kx de nuevos juegos de datos D1 a Dn de otros/nuevos ficheros de registro P1 a Pn.



Para el procesamiento ulterior del contenido del código numérico Kx y/o del código de grupo KGy, KGy+1, estos se decodifican y se emiten como parte alfanumérica del mensaje N1 a Nn del fichero de registro P1 a Pn y pueden almacenarse, por ejemplo, en una memoria 10.

5 Alternativa o adicionalmente pueden estos códigos numéricos Kx decodificados y que representan la parte alfanumérica del mensaje N1 a Nn y los códigos de grupo KGy, KGy + 1 se alimentan a la unidad deflectora 4 para la emisión sobre una pantalla y para los ulteriores análisis y evaluación del/de los evento/s.

10 Por otra parte, el procedimiento permite, que al iniciar el procedimiento sólo se emplee un número predeterminado de grupos Gy y no se generen nuevos grupos Gy+1. De este modo se posibilita primero una agrupación robusta de nuevos códigos numéricos a considerar Kx. En la operación continua puede permitirse entonces la generación de nuevos grupos Gy+1.

Aunque la invención se ha ilustrado y descrito con más detalle mediante ejemplos de ejecución preferidos, la invención no está limitada por los ejemplos mostrados y el experto puede deducir otras variaciones, sin abandonar el alcance de la invención. Particularmente, el compresor de registros 9 puede implementarse en una posición de un dígito apropiada en un componente del sistema de automatización 1.

15

**REIVINDICACIONES**

1. Procedimiento para el procesamiento automático de un número de ficheros de registro (P1 a Pn) de diferente tipo de un sistema de automatización (1) de una instalación técnica, comprendiendo los siguientes pasos:
  - 5           - Determinación de al menos una parte del mensaje (N1 a Nn) de cada juego de datos (D1 a Dn) de un fichero de registro (P1 a Pn) mediante
  - Concatenación de todos los contenidos del respectivo juego de datos (D1 a Dn) del fichero de registro (P1 a Pn),
  - Reunión de los juegos de datos (D1 a Dn) de todos los ficheros de registro (P1 a Pn) en un archivo resumen (SD),
  - 10          - Clasificación cronológica de los juegos de datos (D1 a Dn) de todos los ficheros de registro (P1 a Pn) en el archivo resumen (SD),
  - Compresión de cada juego de datos (D1 a Dn) en el archivo resumen (SD),
  - Codificación de los juegos de datos (D1 a Dn) comprimidos en el archivo resumen (SD) mediante al menos uno o varios códigos numéricos, alfabéticos o alfanuméricos (Kx),
  - 15          - Ordenación y asignación de los juegos de datos comprimidos (D1 a Dn), en base al /a los códigos numéricos, alfabéticos o alfanuméricos (Kx) determinados, a al menos un grupo (Gy) con un código de grupo (KGy) asociado, y
  - decodificación del código de grupo (KGy) del grupo asignado (Gy) y emisión del código de grupo decodificado (KGy) como parte alfanumérica del mensaje (N1 a Nn) del fichero de registro (P1 a Pn), y
  - 20          - Almacenamiento de la parte alfanumérica del mensaje (N1 a Nn) del fichero de registro (P1 a Pn) en una memoria (10).
2. Procedimiento según la reivindicación 1, donde, al comprimir los juegos de datos (D1 a Dn), estos se procesan de tal forma que se reduzca(n) al menos la longitud de los datos y/o el contenido de los datos del respectivo juego de datos (D1 a Dn).
- 25   3. Procedimiento según la reivindicación 2, donde se insertan automáticamente espacios para los caracteres especiales contenidos en un juego de datos (D1 a Dn) y para al menos dos o más espacios consecutivos resultantes de esto se elimina del juego de datos (D1 a Dn) un número de espacios consecutivos más allá de un único espacio.
4. Procedimiento según una de las anteriores reivindicaciones, donde la concatenación de los contenidos del respectivo juego de datos (D1 a Dn) se resuelve y los contenidos alfanuméricos separados resultantes con menos de cuatro caracteres se eliminan.
- 30   5. Procedimiento según la reivindicación 4, donde los contenidos alfanuméricos separados se analizan y unifican de tal forma, que a cada contenido alfanumérico separado se le asigna una palabra de referencia (Rx) y a cada palabra de referencia (Rx), un código numérico (Kx).
- 35   6. Procedimiento según la reivindicación 5, donde los datos de suma (SD) se convierten con los contenidos alfanuméricos separados en base a las determinadas palabras de referencia (Rx) y/o del/de los código(s) numérico(s) (Kx) en una matriz de códigos bidimensionales, particularmente alfanuméricos (KM).
7. Procedimiento según la reivindicación 6, donde los códigos numéricos (Kx) de la matriz de códigos (KM) bidimensional se ordenan cronológicamente y/o en función de la respectiva longitud de caracteres.
- 40   8. Procedimiento según la reivindicación 6 ó 7, donde los códigos numéricos (Kx) similares de la matriz de códigos (KM) se agrupan por medio de una operación de similitud, particularmente por la llamada distancia de Levenshtein.
9. Procedimiento según la reivindicación 8, donde para una distancia de Levenshtein de cero para un código numérico a considerar (Kx) respecto a uno de los grupos (Gy), este código numérico a considerar (Kx) se asigna a este grupo (Gy).

10. Procedimiento según la reivindicación 8 ó 9, donde para una distancia de Levenshtein de uno para un código numérico a considerar (Kx) respecto a uno de los grupos (Gy) y para una diferencia del código numérico a considerar (Kx) sólo en una posición de un dígito de al menos uno de los grupos (Gy), este código numérico a considerar (Kx) se asigna a por lo menos un grupo (Gy).
- 5 11. Procedimiento según una de las anteriores reivindicaciones, donde cada nuevo código numérico a considerar (Kx) se compara con códigos numéricos no-agrupados (Kx) de la misma longitud o con una mayor longitud de como máximo un carácter.
- 10 12. Procedimiento según la reivindicación 11, donde, cuando la distancia de Levenshtein de un código numérico a considerar (Kx) respecto a otro código numérico no-agrupado (Kx) sea igual a uno, entonces se forma un nuevo grupo (Gy+1) de ambos códigos numéricos (Kx) de tal forma que
- para una longitud diferente de ambos códigos numéricos a considerar (Kx), al nuevo grupo (Gy+1) se le asigna el código numérico más corto (Kx) o
  - para una misma longitud de ambos códigos numéricos (Kx) con al menos un carácter diferente, al nuevo grupo (Gy+1) se le asigna el código numérico (Kx) sin el carácter diferente.
- 15 13. Procedimiento según una de las anteriores reivindicaciones, donde, cuando se identifique un código numérico a considerar (Kx), que no presente una distancia de Levenshtein de igual a uno respecto a otro código numérico no-agrupado (Kx), este código numérico a considerar (Kx) se asigna a un nuevo grupo (Gy+1).
14. Empleo del procedimiento según una de las reivindicaciones 1 a 13 en un sistema de automatización (1) de una instalación técnica.
- 20 15. Sistema de automatización (1) de una instalación técnica, comprendiendo al menos un nivel de operación y de monitorización (BB), un nivel de automatización (AS) y un nivel de campo (FE), que están conectados entre sí a través de un sistema de comunicaciones, donde en al menos uno de los niveles se implementa el procedimiento según una de las reivindicaciones 1 a 13.

25

FIG 1

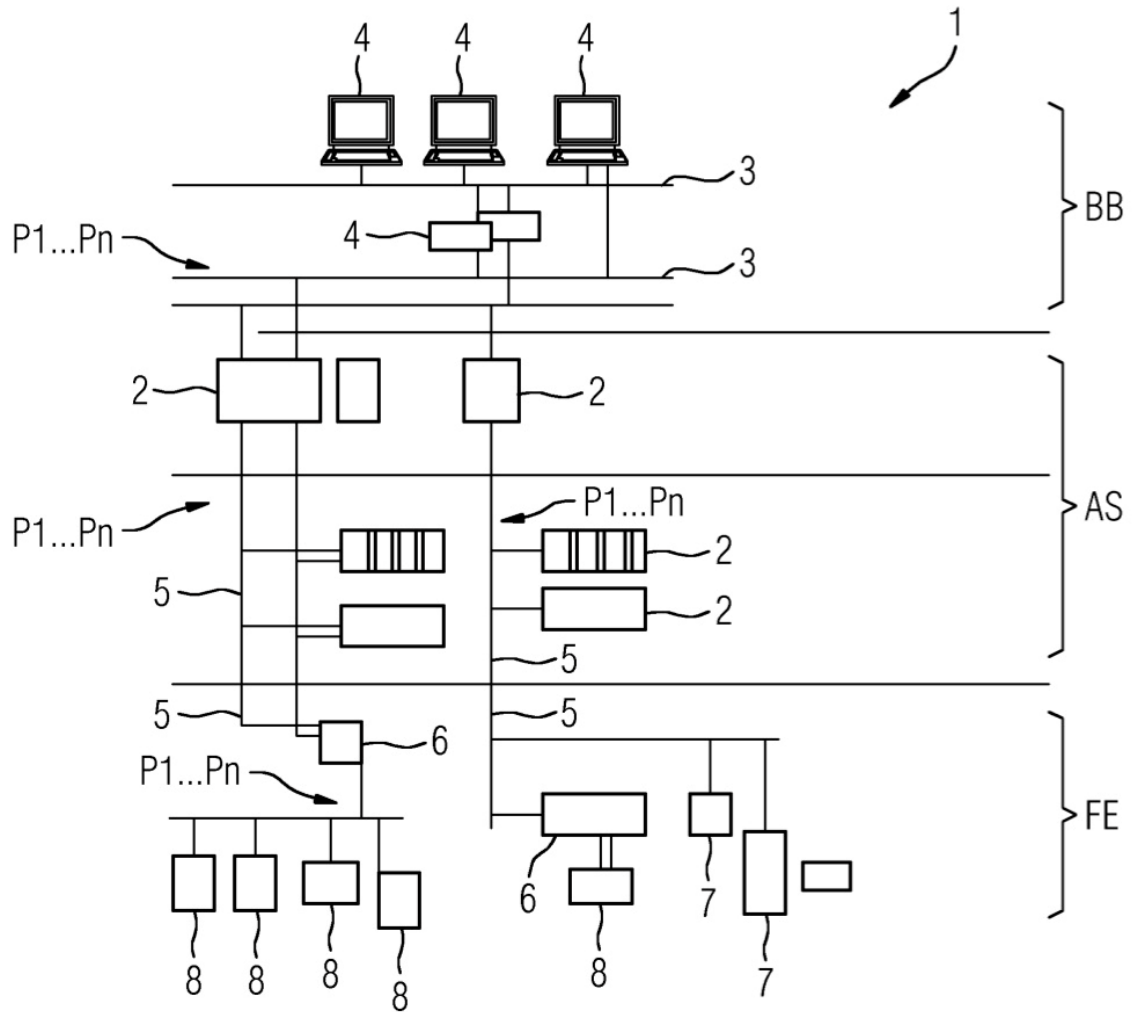


FIG 2

