(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 708 682**

(51) Int. CI.:

H04L 29/06 (2006.01)
H04M 3/38 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Fecha de presentación y número de la solicitud europea: **05.12.2012 E 12195697 (3)**

(97) Fecha y número de publicación de la concesión europea: **31.10.2018 EP 2602982**

(54) Título: **Autentificación de participantes en un servicio de telefonía**

(30) Prioridad:

05.12.2011 DE 102011056038

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
10.04.2019

(73) Titular/es:

AUTHADA GMBH (100.0%)
Julius-Reiber-Strasse 15 a
64293 Darmstadt, DE

(72) Inventor/es:

WIENS, TORSTEN;
PLIES, ANDREAS y
MASSOTH, MICHAEL, PROF. DR.

(74) Agente/Representante:

LINAGE GONZÁLEZ, Rafael

DESCRIPCIÓN

Autentificación de participantes en un servicio de telefonía

La presente invención se refiere a la autentificación de participantes en un servicio de telefonía y, en particular, a un procedimiento, un sistema y un dispositivo que permitan la autentificación de los distintos participantes.

En aplicaciones de seguridad crítica, los interlocutores de las conversaciones telefónicas tienen que confiar en la autenticidad de sus colocutores. Esto afecta, por ejemplo, a la comunicación con personas de determinados colectivos profesionales, que están sujetas a confidencialidad por motivos de ética profesional, o a la comunicación con agencias y otros proveedores de servicios, como los bancos y otros servicios de información, en la que es necesario emitir información crítica para la seguridad durante la comunicación.

En este tipo de aplicaciones, generalmente se verifica inicialmente la supuesta identidad de una de las partes o de uno de los participantes en la comunicación, por ejemplo, mediante el intercambio de información de autentificación. En estos casos, la autenticidad de un participante se entiende generalmente como la autenticidad y credibilidad del participante, que puede verificarse en base a una identidad inequívoca con propiedades características. Estas se verifican mediante procedimientos de autentificación adecuados, en los que es necesario demostrar que la supuesta identidad del participante se corresponde con sus propiedades características.

No obstante, un problema fundamental de este procedimiento consiste en que la identidad de uno de los participantes puede cambiar durante la comunicación. Esto puede ocurrir, por ejemplo, debido a un intento de suplantación de identidad durante una sesión de comunicación existente, en el que un atacante adopta la identidad del participante y actúa en su nombre. Además, existe el problema de una identidad falsamente reivindicada que, no obstante, el procedimiento de autentificación detecta incorrectamente como verdadera. En estos casos, el atacante o bien conoce la información de autentificación del usuario, o explota las debilidades del procedimiento de autentificación. Otro problema de los procedimientos de autentificación consiste en que los datos que se transmiten durante la autentificación pueden ser leídos por un atacante, lo que permite a dicho atacante autentificarse posteriormente como el participante real con los datos interceptados.

Para resolver estos problemas, se conocen enfoques en los que no se produce una transferencia directa de la información de autentificación, de modo que ya no es el propio participante el que emite los datos necesarios para la autentificación, por ejemplo, en los procedimientos de conocimiento cero o de desafío-respuesta. En estos sistemas, el participante solo tiene que aportar una prueba de que posee los datos. También se conocen soluciones que utilizan los datos necesarios para la autentificación una sola vez o que emiten la información de autentificación por un canal independiente, el denominado segundo canal. Sin embargo, incluso con estos enfoques sigue existiendo el problema de que no siempre es posible asignar de forma inequívoca la información de autentificación utilizada a la identidad real de una persona y la autentificación puede estar sujeta a otras restricciones.

Asimismo, se conocen procedimientos de autentificación basados en el reconocimiento vocal o biométrico del hablante, que pueden ser textuales o independientes del texto. Este tipo de procedimientos se conocen particularmente por situaciones de banca telefónica y en ocasiones, se combinan con la transmisión de información de autentificación convencional, como el PIN/TAN. No obstante, estos procedimientos de autentificación requieren formación específica sobre las características vocales de los respectivos participantes para poder afirmar de forma fiable la autenticidad de los participantes. Por lo tanto, debido a la formación necesaria, estos procedimientos exigen más tiempo para la autentificación y son propensos a fallar en caso de posibles cambios o fallos en las características de emisión del canal o a cambios en la voz del usuario, por ejemplo, debido enfermedades u otras circunstancias.

El documento WO 2007/131523 A1describe un procedimiento de autentificación que puede utilizarse para autentificar mensajes emitidos como paquetes en una red de comunicaciones entre un remitente y un receptor. Para ello, está previsto un módulo emisor de control que almacena los datos emitidos y genera datos de autentificación a partir de los mismos, que posteriormente, se transmiten a un módulo receptor de control por un canal de datos seguro. A su vez, el módulo receptor de control almacena los datos recibidos a través del canal y calcula los datos de autentificación correspondientes a partir de los mismos. La autenticidad del mensaje puede determinarse comparando los datos de autentificación.

El documento US 2005/0063522 A1describe un sistema de autentificación biométrico que procesa muestras de voz. En una primera etapa, se captura y almacena una muestra de voz auténtica, que a continuación, un sistema de comparación utiliza para compararla con las muestras de voz introducidas.

El documento WO 03/075540 A2describe un enfoque de autentificación que emplea varios factores, uno de los cuales es un patrón de habla del usuario. En este caso, los patrones de habla se envían a un subsistema de verificación de hablantes en la red de comunicación y se utilizan para la autentificación.

Por lo tanto, un objetivo de la presente invención es resolver los problemas anteriormente mencionados y, en particular, proporcionar un enfoque que permita una autentificación fácil y rápida de los participantes en un servicio de telefonía que garantice la integridad de la comunicación y la autenticidad de los participantes durante la comunicación.

Este objetivo se consigue mediante un procedimiento de autentificación según la reivindicación principal, así como mediante un sistema y un dispositivo para la autentificación de participantes, tal y como se definen en las reivindicaciones dependientes.

El procedimiento para la autentificación de participantes en un servicio de telefonía según la invención comprende el establecimiento de un canal de transmisión entre un primer y al menos un segundo terminal de los participantes en el servicio de telefonía, en el que las señales de voz que se transmiten a través del canal de transmisión se registran en cada terminal y en cada terminal se genera un conjunto de datos de verificación a partir de las señales de voz registradas. Además, se envía al menos un primer conjunto de datos de verificación de los conjuntos de datos prueba disponibles desde el primer terminal por un canal de seguridad y el al menos un primer conjunto de datos de verificación enviado a través del canal de seguridad se compara con al menos un segundo conjunto de datos de verificación de los conjuntos de datos de verificación disponibles.

Mediante el registro sincronizado de las señales de voz y la generación del conjunto de datos de verificación a partir de las señales de voz registradas en cada terminal, se genera una referencia para la señal de voz transmitida a través del canal de transmisión en el terminal correspondiente. En este caso, la señal de voz transmitida a través del canal de transmisión y registrada en uno de los terminales puede comprender las señales de voz enviadas por el terminal, las señales de voz recibidas por el terminal y, preferentemente, tanto las señales de voz enviadas como las recibidas, de modo que las señales de voz registradas en el primer terminal, a partir de las cuales se genera al menos un primer conjunto de datos de verificación, coinciden esencialmente con las señales de voz registradas en el segundo terminal, a partir de las cuales se genera el segundo conjunto de datos de verificación. Al enviar al menos uno de los conjuntos de datos de verificación, por ejemplo, el primer conjunto de datos de verificación, como conjunto de datos de verificación remoto a través del canal de seguridad y compararlo con al menos un conjunto de datos de verificación local coincidente en el tiempo y proveniente de otro terminal, por ejemplo, con el segundo conjunto de datos de verificación, se verifica la autenticidad de al menos uno de los participantes durante la comunicación y, por lo tanto, se verifica, al menos en un lado de la comunicación, la integridad del canal de transmisión durante el tiempo que dure el registro. Así, un participante remoto (y por tanto, también, al menos en parte, el canal de transmisión) se considerará auténtico para una participante local si la comparación del conjunto de datos de verificación enviado con un conjunto de datos de verificación local correspondiente es satisfactoria, de lo que se deduce que al menos las señales de voz transmitidas a través del canal de transmisión por el participante remoto coinciden esencialmente con las recibidas por el participante local. De este modo, la autenticidad de los participantes y la integridad del canal de transmisión pueden garantizarse mediante una repetición del procedimiento y mediante un registro sincronizado y continuado correspondiente en todos los terminales.

En este caso, la comunicación y el servicio de telefonía no se limitan a una comunicación telefónica clásica entre dos terminales de dos participantes. Por el contrario, distintos terminales y una pluralidad de participantes pueden participar en una telecomunicación, por ejemplo, en forma de conferencia telefónica. Los terminales pueden ser cualquier terminal apto para la comunicación de voz, por ejemplo, teléfonos clásicos como teléfonos fijos, teléfonos móviles, teléfonos inalámbricos o teléfonos por satélite, y otros terminales telefónicos, por ejemplo, teléfonos inteligentes, pads y los dispositivos informáticos que puedan utilizarse para la comunicación entre los participantes en un servicio de telefonía. Asimismo, pueden asignarse varios terminales conectados entre sí a un único participante. Además, un terminal puede comprender varios dispositivos que se comunican entre sí, por ejemplo, un dispositivo telefónico acoplado a un servidor de comunicaciones. En este caso, tanto el teléfono como el servidor de comunicaciones o uno de los componentes pueden ser el terminal, de modo que, por ejemplo, las señales de voz que se transmiten a través del canal de transmisión son registradas por el servidor de comunicaciones, que, a su vez, puede generar un conjunto de datos de verificación a partir de las señales de voz registradas. El participante puede ser una persona real o un participante virtual, por ejemplo, un sistema de voz y comunicaciones proporcionado por un ordenador, que se comunica con el otro participante mediante el servidor de comunicaciones. Junto a la transmisión de voz clásica mediante el servicio de telefonía, o además de ésta, también puede proporcionarse una transmisión de audio/vídeo, una transmisión de voz e imagen y una transmisión de datos, por

ejemplo, de datos bancarios. El resto de datos y señales transmitidos pueden utilizarse, además de las señales de voz registradas, para generar los conjuntos de datos de verificación correspondientes para la autentificación en los terminales de los respectivos participantes.

Para registrar las señales de voz, pueden utilizarse interfaces existentes de los terminales correspondientes que ya se utilicen para la entrada de voz de los participantes, por ejemplo, un micrófono integrado en el terminal u otro componente o interfaz de hardware. Además, dicho componente o interfaz también puede implementarse como un software capaz de comunicarse con una interfaz de comunicación correspondiente del terminal y de registrar las señales de voz. Lógicamente, también puede utilizarse para el registro un dispositivo adicional, por ejemplo, un micrófono externo, que puede conectarse o bien, directamente al terminal o suministrarse en forma de auriculares o de micrófono de sala para registrar las señales de voz del participante.

El procedimiento según la invención permite evitar una pluralidad de ataques al servicio de telefonía, por ejemplo, intentos de suplantación de identidad y otros intentos de apoderarse del canal de transmisión o de adoptar la identidad de uno de los participantes a fin de interferir en la integridad de la comunicación. Gracias el enfoque según la invención, los posibles atacantes solo podrían simular los conjuntos de datos de verificación coincidentes de los respectivos participantes e infiltrarlos en el canal de seguridad con un esfuerzo técnico muy elevado, e incluso es posible que no resultase factible en la práctica.

Por lo tanto, el procedimiento según la invención ofrece un grado de seguridad muy elevado y es de aplicación universal, ya que no depende en particular de un determinado terminal ni sistema o red telefónicos. Además, el procedimiento también es robusto, ya que todos los participantes pueden reaccionar a posibles cambios en la señal de transmisión gracias al registro de las señales de voz en cada terminal. Otras interferencias de comunicación pueden suprimirse eficazmente mediante procedimientos de procesamiento y análisis de señales adecuados al generar los conjuntos de datos de verificación en los terminales. Además, el procedimiento según la invención no requiere recursos informáticos excesivos, por lo que puede utilizarse desde ya en terminales móviles, por ejemplo, teléfonos inteligentes, pads o tabletas. Por último, el procedimiento según la invención no exige adaptar el procedimiento de autentificación a cada uno de los participantes, por ejemplo mediante una formación, por lo que puede utilizarse de forma fiable y flexible para la comunicación oral entre unos participantes cualquiera.

En una realización preferida de la presente invención, el procedimiento comprende, además, la recepción del primer conjunto de datos de verificación por el segundo terminal a través del canal de seguridad y la comparación del primer conjunto de datos de verificación recibido a través del canal de seguridad con el segundo conjunto de datos de verificación mediante el segundo terminal. De este modo, el participante del segundo terminal puede verificar la autenticidad del participante del primer terminal, permitiendo una comprobación unilateral de la integridad del servicio de telefonía. Por ejemplo, un proveedor del servicio de telefonía puede garantizar que la identidad de los participantes, por ejemplo, sus clientes, no cambia en caso de que un tercero se apropie de forma no deseada el canal de transmisión entre el proveedor y un cliente. Para ello, las señales de voz pueden registrarse en los terminales de los clientes y los conjuntos de datos de verificación correspondientes pueden enviarse a través del canal de seguridad al proveedor, que en este caso se comunica mediante el segundo terminal, por ejemplo, un servidor de comunicaciones, que, a su vez, puede compararlos con su segundo conjunto de datos de verificación local, por ejemplo, en el servidor de comunicaciones. La autenticidad del proveedor también podría verificarse en los terminales de los clientes para garantizar que el cliente no se comunica con un tercero que se haya apropiado del canal de transmisión. Para ello, el proveedor puede registrar la señal de voz en su terminal, en este caso, el primer terminal, y enviarla a sus clientes a través del canal de seguridad, que a su vez, compararán el conjunto de datos de verificación con su conjunto de datos de verificación local. De este modo, puede realizarse al menos una verificación unilateral de la integridad del canal de transmisión y, por lo tanto, una autentificación unilateral.

No obstante, preferentemente, el segundo conjunto de datos de verificación también se envía desde el segundo terminal a través del canal de seguridad. Para ello, el procedimiento comprende, además, la recepción del segundo conjunto de datos de verificación por el primer terminal a través del canal de seguridad y la comparación del segundo conjunto de datos de verificación recibido a través del canal de seguridad con el primer conjunto de datos de verificación mediante el primer terminal. El intercambio recíproco de los conjuntos de datos de verificación abre la posibilidad de que, además de la autenticidad del participante del primer terminal, también pueda verificarse de forma recíproca e independiente, la autenticidad del participante del segundo terminal. Por lo tanto, cada participante adicional del servicio de telefonía puede, a su vez, enviar su conjunto de datos de verificación correspondiente a través del canal de seguridad, lo que permite verificar su autenticidad. De este modo, se consigue de forma particularmente eficaz una verificación bilateral o multilateral de los participantes en el servicio de telefonía. En particular, cada uno de los participantes puede, además, enviar su conjunto de datos de verificación a través del canal de seguridad y/o recibir el resto de conjuntos de datos de verificación a través del canal de seguridad, así

como compararlos con la copia local de su conjunto de datos de verificación. La autentificación recíproca y, en gran medida, independiente permite alcanzar un grado de seguridad elevado, lo que hace prácticamente imposible que un atacante suplante las identidades de los participantes o ponga en peligro el servicio de telefonía.

En una realización preferida de la presente invención, el procedimiento comprende, además, el envío del segundo conjunto de datos de verificación desde el segundo terminal a través del canal de seguridad y la comparación del primer conjunto de datos de verificación recibido a través del canal de seguridad con el segundo conjunto de datos de verificación enviado a través del canal de seguridad mediante una unidad de comparación acoplada al canal de seguridad que recibe los dos conjuntos de datos de verificación enviados. Gracias a esto, no es necesario realizar la comparación en cualquier terminal o en un solo terminal que, según el caso, puede disponer de recursos informáticos limitados. Por el contrario, los procedimientos de comparación pueden ejecutarse en una unidad de comparación configurada según corresponda, por ejemplo, en un ordenador que forme parte de una infraestructura de seguridad y que tenga una potencia mayor que los terminales. También este enfoque proporciona una forma particularmente segura de verificar la autenticidad de los respectivos participantes en el servicio de telefonía, que puede ejecutarse mediante una unidad de comparación centralizada y fiable. Puede informarse a los distintos terminales del resultado de la verificación mediante mensajes adecuados a través del canal de seguridad. De forma alternativa o adicional, la unidad de comparación puede influir, además, en la comunicación a través del canal de transmisión y, si la probabilidad de una suplantación no deseada de la identidad de uno de los participantes es lo suficientemente elevada, apropiarse, redirigir, volver a establecer o interrumpir el canal de transmisión.

En una realización particularmente preferida, el procedimiento comprende, además, la fijación de un momento de inicio del registro de las señales de voz por uno de los terminales, registrándose las señales de voz a partir del momento de inicio en los terminales, esencialmente en el mismo momento o en un momento posterior, pospuesto por un valor de demora. Para ello, preferentemente se sincronizan los finales de los participantes, por ejemplo, utilizando el protocolo Network Time Protocol (NTP, por sus siglas en inglés). Un terminal recibe la función de un servidor, por ejemplo, cuando los terminales negocian las funciones entre sí, o cuando un proveedor de servicios asigna la función de servidor a su terminal y a los terminales de los participantes, el papel de cliente. Ahora, el servidor puede definir el momento de inicio y emitir un momento t_0 correspondiente a todos los demás terminales. Aunque el servidor puede fijar libremente el momento t_0, este debe coincidir ventajosamente con el inicio de la llamada. No obstante, el momento t_0 también puede seleccionarse en función de otros parámetros, factores o circunstancias del sistema, por ejemplo, el tiempo de respuesta de un servicio web o del canal de transmisión, o en base a consideraciones generales relacionadas con el momento en que se desea verificar la autenticidad de los participantes, por ejemplo, cuando vayan a emitirse datos críticos para la seguridad durante la comunicación. Además, puede preverse una compensación de las posibles demoras en el canal de transmisión para lograr una sincronización lo más precisa posible de los participantes.

Los terminales pueden empezar a registrar las señales de voz correspondientes inmediatamente en el momento t_0 acordado. Sin embargo, dado que pueden producirse demoras en la transmisión a través del canal de transmisión, las señales de voz también pueden registrarse en distintos momentos que difieren entre sí en un valor de demora. Este valor de demora puede derivarse de la demora de transmisión real entre el servidor y un terminal. Por ejemplo, en primer lugar, el servidor puede medir un tiempo de ida y vuelta (RTT, por sus siglas en inglés) para determinar cuánto tiempo tarda una señal en recorrer la distancia definida por el canal de transmisión desde el servidor hasta el terminal y viceversa. Esta medición puede repetirse varias veces y, a partir del conjunto de tiempos de RTT obtenidos, por ejemplo, a partir de la media, la mediana o cualquier otra estimación adecuada, el servidor puede fijar un valor de demora para la transmisión, por ejemplo, un tiempo en un solo sentido (STT, por sus siglas en inglés) según $RTT = 2 \cdot STT$, e informar al terminal del valor del demora. A continuación, el servidor puede iniciar el registro de las señales de voz en el momento t_0 y el terminal cliente puede iniciarlo en el momento $t_0 \pm STT$. La decisión de demorar o adelantar el registro puede depender de qué señales de voz provenientes de qué fuente vayan a desempeñar un papel importante en la verificación de la autenticidad. Si, por ejemplo, las señales de voz del servidor son las decisivas, el cliente puede iniciar el registro en el momento $t_0 + STT$. No obstante, el servidor también puede indicar al terminal cliente que inicie el registro en un momento $t_0 - STT$ si la autentificación va a basarse principalmente en señales de voz emitidas desde el terminal al servidor. El servidor puede fijar un valor de demora STT_i para cada uno de los terminales y comunicarse con ese terminal, o puede calcular un promedio a partir de la medición de los RTTs de todos los terminales y derivar un STT promedio a partir de los mismos, que se utilizará como valor de demora de la recepción en cada terminal.

En una realización particularmente preferida de la presente invención, el registro de las señales de voz comprende el registro de un segmento de la señal de voz dividida en un conjunto de ventanas. Estas ventanas definen un segmento de señal más pequeño y comprenden la señal de voz correspondiente. De este modo, el registro del segmento de la señal de voz puede iniciarse en el momento de inicio t_0 predefinido (o en un momento $t_0 \pm STT_i$ con la

demora correspondiente), por lo que el segmento de la señal de voz puede tener una longitud de, por ejemplo, Δt_m. El segmento de la señal de voz Δt_m puede subdividirse en un conjunto de n ventanas F_i con una longitud de Δt_F, que pueden ser directamente adyacentes o solaparse dependiendo de un factor de solapamiento O_F, en el que se aplica $0 \leq O_F \leq 1$. En este caso, la duración del solapamiento puede ser $O_F \cdot \Delta t_F$. El uso de distintas ventanas permite analizar y comparar específicamente la señal de voz, mejorando la calidad del procesamiento posterior.

En otra realización de la presente invención, cada conjunto de datos de verificación comprende al menos una de las ventanas. Por ejemplo, la señal de voz registrada definida por la ventana puede copiarse en el conjunto de datos de verificación. Los conjuntos de datos de verificación pueden comprender varias ventanas o todo el conjunto de ventanas. Las señales de voz vinculadas a las respectivas ventanas también pueden someterse a una filtración y procesamiento previo adecuados utilizando, por ejemplo, diversos filtros de ventana, de interferencia o de alisado.

En una realización particularmente preferida, la generación del conjunto de datos de verificación comprende un cálculo de las características de voz de al menos una de las ventanas y la incorporación de las características de voz al conjunto de datos de verificación. Las características de voz pueden calcularse utilizando cualquier procedimiento de extracción de características de audio. Para ello, la señal puede someterse a un procesamiento previo. De este modo, cada ventana puede normalizarse volviendo a calcular las amplitudes individuales en función de la amplitud local máxima de una ventana y de la amplitud máxima permitida o escalarlas con un factor adecuado. Además, las amplitudes de las frecuencias bajas y altas pueden equiparse mediante un simple filtro digital (procedimiento de preénfasis). La señal de la ventana en cuestión también puede filtrarse mediante una función de ventana, por ejemplo, mediante la función de ventana de Hamming.

Para extraer las características reales, la señal que se desea analizar puede transformarse en un espacio de frecuencias, por ejemplo, aplicando la transformada rápida de Fourier (FFT, por sus siglas en inglés). A continuación, las características que caracterizan la señal de voz en la ventana en cuestión pueden derivarse a partir de los coeficientes de Fourier. Por ejemplo, pueden calcularse una medida de la componente de ruido de la señal (compacidad), una medida del cambio de espectro de la señal (flujo espectral), una desviación estadística estándar del espectro de magnitud (variabilidad espectral), otras características estadísticas tales, como momentos individuales de una distribución estadística, y otras características. Asimismo, pueden combinarse otros procedimientos de filtrado y extracción de características para obtener características significativas para la señal de voz subyacente, por ejemplo, la escala Mel, el cepstrum y los coeficientes cepstrales en las frecuencias de Mel (MFCC, por sus siglas en inglés) derivados de los mismos. Además, los filtros individuales pueden parametrizarse y organizarse adecuadamente en un banco de filtros para derivar las características a partir de los datos de la señal de voz. Para ello, los parámetros de los filtros pueden ser predeterminados, o puede definirlos determinísticamente, por ejemplo, el servidor, o puede realizarse una parametrización determinística aleatoria si el servidor distribuye un valor de inicio para los generadores de números pseudoaleatorios, que generan parámetros aleatoriamente, pero generando el mismo en todos los terminales, para los distintos filtros y etapas de procesamiento a partir del valor de inicio.

El cálculo de los coeficientes cepstrales en las frecuencias de Mel, por ejemplo, como algoritmo MFCC, puede comprender las siguientes etapas de cálculo: 1. Creación de una ventana de Hamming y realización de otras etapas de procesamiento previo de la señal, 2. Cálculo de la FFT, 3. Creación del espectro de magnitud, 4. Aplicación de un banco de filtros, 5. Cálculo de logaritmo de los valores del banco de filtros y 6. Decorrelación con una transformada de coseno discreta.

Según la invención, pueden calcularse tanto características de voz individuales como un grupo de características de voz para una ventana, que posteriormente se incorporan en su totalidad al conjunto de datos de verificación y pueden utilizarse para la comparación.

Preferentemente, la comparación comprende la definición de una selección determinística aleatoria del conjunto de ventanas y la definición de una distancia entre los datos vinculados a las ventanas seleccionadas a partir del primer conjunto de datos de verificación enviado a través del canal de seguridad y el segundo conjunto de datos de verificación. La definición determinística aleatoria de las distintas ventanas utilizadas para la comparación hace que el procedimiento sea más robusto en general y dificulta los ataques contra el procedimiento de autentificación. Para ello, las distintas ventanas o los índices de tiempo relativos a las ventanas que se desea analizar F_i pueden definirse de forma determinística aleatoria una vez que se hayan registrado los segmentos de la señal de voz correspondientes Δt_m y las ventanas asociadas F_i en todos los terminales. Para este fin, puede tomarse el valor de inicio distribuido inicialmente por el servidor u otro valor de inicio, que puede ser utilizado por cualquier terminal para inicializar un generador de números pseudoaleatorios, lo que permite calcular nuevos números pseudoaleatorios. Por lo tanto, estos números son iguales y conocidos por todos los participantes, pero no por terceros, en particular,

por atacantes que no dispongan información sobre el valor de inicio.

Preferentemente, la selección de ventanas puede dividirse en dos grupos, pudiendo disponerse en el primer grupo ventanas con un mismo índice de tiempo y que, por lo tanto, corresponden al mismo (o pospuesto por un valor de demora) segmento temporal de la señal de voz. En el segundo grupo, pueden disponerse ventanas adicionales con distintos índices de tiempo y que, por lo tanto, corresponden a distintos segmentos temporales. Al comparar los conjuntos de datos de verificación, pueden verificarse las ventanas del primer grupo para ver si son similares o idénticas y las ventanas del segundo grupo para ver si son distintas. Esta consideración particularmente ventajosa de las ventanas desde distintos momentos aumenta significativamente la calidad de asignación y el índice de detección general durante la autentificación.

La selección determinística aleatoria del conjunto de ventanas puede definirse después de la recepción del primer conjunto de datos de verificación y, en caso necesario, de otros conjuntos de datos de verificación a través del canal de seguridad. En este caso, los conjuntos de datos de verificación pueden contener datos de todas las ventanas del segmento de señal de voz registrado. Alternativamente, la selección también puede definirse antes de que se generen los conjuntos de datos de verificación, de modo que solo los datos correspondientes a las ventanas que se van a comparar pueden incorporarse al conjunto de datos de verificación correspondiente. Aunque un atacante podría obtener algún indicio sobre los parámetros del procedimiento a partir de la transferencia de únicamente una selección de ventanas, lo que teóricamente, podría permitirle superar el procedimiento de autentificación, este posible punto de ataque debe evaluarse en relación con la reducción del ancho de banda necesario del canal de seguridad, que puede aprovecharse de forma más eficaz. Además, también es posible transmitir datos de un conjunto mixto de ventana que contenga tanto datos procedentes de la selección de ventanas como otros datos adicionales incorporados aleatoriamente procedentes de ventanas, de modo que la selección real permanece oculta al atacante, permitiendo, además, reducir el ancho de banda necesario del canal de seguridad. Además, un terminal puede transmitir únicamente los datos procedentes de las ventanas seleccionadas al conjunto de datos de verificación correspondiente y otro terminal puede transmitir los datos procedentes de todo el conjunto de ventanas. En particular, la elección de la tecnología de transmisión adecuada puede depender de la capacidad de transmisión que ofrezca por el canal de seguridad, de modo que, por ejemplo, con una conexión inalámbrica, solo podrá transmitirse un conjunto de datos reducido.

Los índices de tiempo de las ventanas calculados de forma determinística aleatoria se utilizan para identificar los datos correspondientes del conjunto de datos de verificación recibido y verificarlos con una copia del conjunto de datos de verificación local a fin de determinar si concuerdan o discrepan. Los datos correspondientes pueden compararse mediante cualquier enfoque de concordancia de patrones. Para ello, los datos correspondientes a las ventanas se registran como vectores de entrada M y N, en los que puede darse $M = \{M_0, M_1, ..., M_m\}$, $N = \{N_0, N_1..., N_n\}$, en el que $m=n$ o $m \neq n$. Para determinar la similitud o disimilitud de los dos vectores de entrada M, N, puede utilizarse cualquier métrica de distancia que tenga suficientemente en cuenta las propiedades de las características de audio. Así, en el caso más sencillo, puede utilizarse la distancia euclídea o una versión ligeramente modificada de la misma que preferentemente, tenga en cuenta secuencias de características con distintas longitudes. También puede utilizarse una versión ampliada del procedimiento de los vecinos más cercanos como mecanismo de clasificación. Asimismo, pueden utilizarse distintas métricas de distancia, por ejemplo, la distancia Manhattan o distancia de Hamming, que permiten diferenciar mejor los coeficientes con una distancia menor y los coeficientes con una distancia mayor debido a las ponderaciones exponenciales. Preferentemente, puede utilizarse un algoritmo de distorsión temporal dinámica (DTW, por sus siglas en inglés) para comparar vectores de características. La DTW calcula una matriz de distancia para cada uno de los vectores de entrada M, N con una longitud distinta $m \neq n$ que reproduce, para cada entrada (i, j), la distancia entre el elemento M_i número i del vector de entrada M y el elemento N_j número j del vector de entrada N. Como métricas de distancia para la matriz de distancia, pueden utilizarse distintos procedimientos, correspondiendo la mejor coincidencia entre dos vectores M, N a la trayectoria más corta P entre la esquina inferior izquierda (1,1) y la esquina superior en (m, n). La longitud de la trayectoria es la suma de todos los valores de la matriz contenidos en la trayectoria. Resulta evidente que la concordancia de patrones puede realizarse varias veces para cada ventana, dependiendo de las características calculadas.

Los resultados de la concordancia de patrones pueden clasificarse, por ejemplo, utilizando el procedimiento de votación por mayoría ponderada para poder sacar conclusiones sobre la autenticidad del participante y el canal de transmisión. Para ello, pueden compararse los distintos resultados parciales de las características individuales y el resultado global con los valores de umbral. Si un resultado parcial o el resultado global supera o no alcanza el valor de umbral correspondiente, el canal y el participante pueden considerase auténticos o no auténticos.

En una realización preferida de la presente invención, el procedimiento comprende, además, el cifrado del canal de seguridad. Esto permite aumentar aún más la seguridad del sistema. No obstante, también resulta evidente que el

procedimiento en su conjunto no puede depender de un cifrado o una protección determinados del canal de seguridad, siempre que los conjuntos de dato de verificación se intercambien a través de un canal lo suficientemente seguro y fiable como para evitar su manipulación por parte de terceros. En particular, las ventanas de comparación definidas de forma determinística aleatoria y las señales de voz registradas por ambas partes ya de por sí reducen significativamente las posibilidades de ataque en el procedimiento según la invención. El canal de seguridad puede ser cualquier conexión segura, preferentemente basada en paquetes. No obstante, de forma alternativa o adicional al cifrado, el canal de seguridad también puede protegerse o blindarse frente a distintos ataques mediante medidas adicionales.

En otra realización preferida, el procedimiento comprende, además, la determinación de la identidad de al menos un participante mediante la lectura de las características de seguridad a partir de una memoria de datos personalizada para el participante y la transmisión de las características de seguridad a un servidor.

Preferentemente, la lectura comprende la lectura de los datos de un documento de identidad. Para ello, el participante puede disponer de un documento de identidad electrónico y de un lector correspondiente con una funcionalidad adecuada, por ejemplo, un lector de nPA y una aplicación AusweisApp (aplicación de documento de identidad), o un teléfono inteligente u otro terminal móvil apto para tecnología NFC (Near Field Communication) capaz de comunicarse directamente con un chip RFID del documento de identidad electrónico para leer los datos y autentificar al participante mediante una infraestructura de seguridad. A tal efecto, el participante puede autentificarse en primer lugar mediante la lectura correspondiente de una identidad electrónica (IDe) del documento de identidad electrónico. Tras la autentificación inicial, según la invención la autenticidad se verifica y constata de forma continua mediante el procedimiento según la invención. Para ello, puede proporcionarse una autentificación de dos factores, en la que pueden combinarse la identidad electrónica y la comparación de las señales de voz mediante los conjuntos de datos de verificación para la autentificación. En particular, en primer lugar, la identidad del participante puede constatase mediante el documento de identidad electrónico y una infraestructura de seguridad correspondiente y, a continuación, garantizarse, preferentemente con los mismos componentes de software, mediante la supervisión continua de la llamada, que no cambie la identidad del participante. De este modo, la infraestructura de seguridad, que puede utilizarse para la verificación inicial de la identidad del participante, también puede emplearse para proporcionar el canal de seguridad para el intercambio de conjuntos de datos de verificación entre los participantes. No obstante, también es posible establecer el canal de seguridad y desconectarlo de la infraestructura de seguridad mediante otro medio de transmisión.

En otra realización ventajosa, el procedimiento comprende, además, el cálculo de una marca de agua. Preferentemente, la marca de agua puede calcularse durante la generación del conjunto de datos de verificación. Asimismo, la marca de agua puede incrustarse en el conjunto de datos de verificación. No obstante, la marca de agua también puede calcularse además de las características del conjunto de datos de verificación e incrustarse directamente en la conversación telefónica. Por ejemplo, tras una verificación inicial de la identidad de al menos uno de los participantes, por ejemplo, mediante un documento de identidad electrónico, pueden transmitirse una marca de agua de audio generada de forma continua o un valor de dispersión (hash) robusto desde el terminal del participante, por ejemplo, a través del canal de seguridad o incrustarse en la conversación telefónica para que el terminal del otro participante las lea y verifique de forma continua. Si la marca de agua no está presente o es incorrecta, la conexión puede considerarse no auténtica. La marca de agua puede, por ejemplo, contener un valor determinístico aleatorio que puede ser conocido por ambos participantes gracias a un valor de inicio acordado al principio y que puede volver a calcularse periódicamente. Las marcas de agua pueden calcularse mediante cualquier procedimiento de cálculo de marcas de agua de audio, que pueden parametrizarse para la transmisión de voz por telefonía. No obstante, la presente invención no se limita a una configuración específica de las marcas de agua o del hash.

Según la invención se proporciona, además, un soporte legible por ordenador con comandos almacenados en el mismo que, al ser ejecutado por un sistema informático, ordena al sistema informático que ejecute el procedimiento según la invención. En particular, los distintos terminales pueden leer los comandos almacenados desde el soporte legible por ordenador, de modo que cada terminal se configura para registrar señales de voz transmitidas a través de un canal de transmisión y generar un conjunto de datos de verificación a partir de la señal de voz registrada. Además, una infraestructura del sistema informático puede leer los comandos almacenados en el soporte legible por ordenador, de modo que la infraestructura se configura para establecer el canal de transmisión entre un primer terminal y al menos un segundo terminal de los participantes en el servicio de telefonía y para proporcionar un canal de seguridad a través del cual se transmite al menos un primer conjunto de datos de verificación del primer terminal desde el primer terminal, con lo que la infraestructura en su totalidad se configura para comparar al menos un primer conjunto de datos de verificación enviado a través del canal de seguridad con al menos un segundo conjunto de datos de verificación del segundo terminal.

Según la invención, también se proporciona un sistema para la autentificación de participantes en un servicio de telefonía que comprende un canal de transmisión que conecta un primer terminal y al menos un segundo terminal de los participantes en el servicio de telefonía. Además, en cada terminal están previstos un dispositivo de registro y una unidad de procesamiento, en el que el dispositivo de registro está configurado para registrar, en el terminal correspondiente, las señales de voz transmitidas a través del canal de transmisión y la unidad de procesamiento está configurada para generar un conjunto de datos de verificación correspondiente a partir de las señales de voz registradas. El sistema comprende, además, un canal de seguridad acoplado al primer terminal y a través del cual se envía al menos un primer conjunto de datos de verificación del primer terminal. El sistema comprende, además, una unidad de comparación configurada para comparar al menos un primer conjunto de datos de verificación enviado a través del canal de seguridad con al menos un segundo conjunto de datos de verificación del segundo terminal.

El sistema según la invención permite una autentificación fácil y segura de los participantes en el servicio de telefonía que no exige engorrosas sesiones de formación y que permite autentificar eficazmente a los participantes en el servicio de telefonía durante la transmisión de voz.

Las señales de voz transmitidas a través del canal de transmisión y registradas por los dispositivos de registro dispuestos en los respectivos terminales pueden comprender tanto las señales de voz de entrada como las de salida de la llamada, en la que las señales de voz registradas en el primer terminal, a partir de las cuales se genera el primer conjunto de datos de verificación, se corresponden esencialmente con las señales de voz registradas en el segundo terminal, a partir de las cuales se genera el segundo conjunto de datos de verificación. No obstante, en una realización, también pueden registrarse y evaluarse las señales de voz analógicas de entrada, con lo que se obtiene, de forma análoga en el segundo terminal, el segundo conjunto de datos de verificación a partir de las señales de voz registradas recibidas por el primer terminal. De ese modo, puede generarse al menos un segundo conjunto de datos de verificación a partir de, al menos, las señales de voz enviadas por el segundo participante a través del canal de transmisión y que se corresponden con las señales de voz recibidas por el primer participante, a partir de las cuales puede generarse al menos un primer conjunto de datos de verificación. Las posibles divergencias resultantes entre las respectivas señales de voz registradas de la llamada en el lado del hablante y del receptor, por ejemplo, en el segundo terminal y en el primer terminal, pueden compensarse mediante una construcción adecuada de los conjuntos de datos de verificación. Por ejemplo, pueden definirse adecuadamente los intervalos de tiempo, los valores de umbral de decisión para la verificación de los intervalos de tiempo y las reglas de vinculación de los resultados de la verificación al conjunto de datos de verificación para compensar las divergencias.

En una realización preferida de la presente invención, el segundo terminal está acoplado al canal de seguridad y a la unidad de comparación para recibir el primer conjunto de datos de verificación y comparar el primer conjunto de datos de verificación recibido a través del canal de seguridad con el segundo conjunto de datos de verificación. Para ello, la unidad de comparación puede diseñarse como un módulo de hardware dedicado o como un módulo de software dispuesto en el segundo terminal. En este caso, el segundo terminal puede determinar eficazmente la autenticidad del participante del primer terminal, así como la autenticidad e integridad del canal de transmisión, comparando el primer conjunto de datos de verificación recibido a través del canal de seguridad con el conjunto de datos de verificación local. Por consiguiente, la autenticidad del participante y del canal de transmisión puede depender, al menos en parte, del resultado de la comparación de la unidad de comparación.

En otra realización preferida, el segundo terminal está configurado, además, para transmitir el segundo conjunto de datos de verificación a través del canal de seguridad, y el primer terminal está configurado, además, para recibir el segundo conjunto de datos de verificación a través del canal de seguridad, en el que el sistema comprende, además, una unidad de comparación adicional acoplada al primer terminal para comparar el segundo conjunto de datos de verificación recibido a través del canal de seguridad con el primer conjunto de datos de verificación del primer terminal. De este modo, la autenticidad del participante del segundo terminal también puede verificarse mediante comparaciones adicionales por parte del resto de participantes con sus conjuntos de datos de verificación. Por lo tanto, también es posible enviar cada conjunto de datos de verificación adicional de otros terminales de otros participantes a través del canal de seguridad y compararlos con una copia del conjunto de datos de verificación local correspondientes del resto de terminales. Así, el participante del segundo terminal puede verificar la autenticidad del participante del primer terminal y el participante del primer terminal puede verificar la autenticidad del participante del segundo terminal, es decir, es posible realizar una autentificación mutua de los participantes y del canal de transmisión de forma independiente entre sí. Por lo tanto, la comunicación puede considerarse fiable, rechazarse o incluso abortarse de forma independiente entre sí. También existe la posibilidad de que otros participantes en el servicio de telefonía realicen la autentificación mutua, en la que se verificar la autenticidad de la estación remota correspondiente comparando el conjunto de datos de verificación recibido con la copia del conjunto de datos de verificación local correspondiente.

En otra realización preferida, el segundo terminal también está configurado para enviar el segundo conjunto de datos de verificación a través del canal de seguridad, y la unidad de comparación está acoplada al canal de seguridad y configurada para recibir los dos conjuntos de datos de verificación enviados y para comparar el primer conjunto de datos de verificación enviado a través del canal de seguridad con el segundo conjunto de datos de verificación enviado a través del canal de seguridad. Asimismo, la unidad de comparación puede recibir otros conjuntos de datos de verificación de otros terminales e incluirlos en la comparación. De este modo, la autentificación de todo el sistema puede realizarse mediante una unidad de comparación central que puede estar integrada, por ejemplo, en una infraestructura de seguridad.

En otra realización de la presente invención, uno de los terminales fija un momento de inicio para el registro y los dispositivos de registro registran las señales de voz a partir del momento de inicio en los terminales, esencialmente en el mismo momento o en un momento posterior, pospuesto por un valor de demora. Para sincronizar los distintos terminales, puede utilizarse un servidor de temporización con un protocolo de sincronización adecuado, por ejemplo, NTP.

En otra realización de la presente invención, el dispositivo de registro registra un segmento de señal de voz que se divide en un conjunto de ventanas.

En una realización adicional de la presente invención, cada conjunto de datos de verificación comprende al menos una de las ventanas.

Según otra realización de la presente invención, el conjunto de datos de verificación generado a partir de las señales de voz comprende las características de voz calculadas a partir de al menos una de las ventanas. Para cada ventana, puede calcularse exactamente una característica de voz o pueden calcularse varias características de voz e integrarlas en el conjunto de datos de verificación.

En otra realización ventajosa de la presente invención, la unidad de comparación está configurada, además, para definir una selección determinística aleatoria del conjunto de ventanas y para definir una distancia entre los datos vinculados a las ventanas seleccionadas a partir del primer conjunto de datos de verificación enviado a través del canal de seguridad y el segundo conjunto de datos de verificación. Para ello, los conjuntos de datos de verificación enviados pueden comprender los datos de todas las ventanas, de modo que los datos correspondientes de la ventana en cuestión pueden extraerse de la selección determinística aleatoria mediante los índices correspondientes y tenerse en cuenta para la comparación. No obstante, la unidad de comparación también puede estar acoplada a la unidad de procesamiento del terminal y ordenar a la unidad de procesamiento que genere únicamente los datos de las ventanas que se corresponden con la selección determinística aleatoria de ventanas. Dado que la selección determinística aleatoria produce el mismo resultado, por ejemplo, mediante la distribución de un valor de inicio en cada terminal, en cada terminal se analizan únicamente las ventanas necesarias para la comparación y se envían al conjunto de datos de verificación a través del canal de seguridad. Esto permite reducir el tiempo de procesamiento y aprovechar mejor la capacidad de transmisión del canal de seguridad.

En una realización preferida de la presente invención, el canal de transmisión es un canal telefónico analógico o digital o proporciona una conexión basada en paquetes, preferentemente una conexión de telefonía por internet o una conexión de voz sobre IP. Dado que el enfoque de autentificación según la invención no requiere un canal de transmisión específico para la transmisión de voz, el enfoque también puede utilizarse de forma particularmente ventajosa con infraestructuras telefónicas ya existentes y otros servicios de telefonía sin necesidad de realizar cambios en la infraestructura. Por lo tanto, el enfoque puede utilizarse de forma inmediata tanto para las conexiones telefónicas clásicas, basadas en una centralita, como para las conexiones telefónicas basadas en paquetes.

En otra realización de la presente invención, el canal de seguridad está configurado para la transmisión de datos digitales. Por ejemplo, el canal de seguridad puede ser una conexión basada en paquetes común a través de una red. No obstante, preferentemente, el canal de seguridad está cifrado. De forma alternativa o adicional, pueden preverse otras medidas de seguridad para garantizar que el canal de seguridad sea lo más seguro posible frente a escuchas e interferencias.

En otra realización ventajosa de la presente invención, el sistema comprende, además, al menos un dispositivo de lectura acopado a un terminal correspondiente de un participante que está configurado para leer las características de seguridad de una memoria de datos personalizada para el participante, estando conectado el dispositivo de lectura a un servidor y configurado para enviar las características de seguridad leídas al servidor a fin de determinar la identidad del participante. El servidor puede formar parte de una infraestructura de seguridad configurada para

verificar las identidades electrónicas de los participantes. Dicha infraestructura de seguridad puede estar proporcionada por un organismo de confianza. Por lo tanto, el canal de seguridad también puede estar proporcionado por la infraestructura de seguridad y coincidir con el canal de comunicación mediante el cual el dispositivo de lectura del terminal se comunica con el servidor para verificar la identidad. No obstante, el canal de seguridad también puede estar desconectado de este canal de comunicación, pudiendo estar diseñado, también, como un canal completamente desconectado de la infraestructura de seguridad.

Preferentemente, la memoria de datos está dispuesta en un documento de identidad y el dispositivo de lectura está configurado para leer los datos de dicha tarjeta de identidad. No obstante, es comprensible que la memoria de datos pueda proporcionarse de otra manera, por ejemplo, en forma de tarjetas de firma o de un chip específico emitido por un organismo de confianza para permitir la verificación de la identidad del participante.

En otra realización, la unidad de procesamiento está configurada, además, para calcular una marca de agua.

Según la invención se proporciona, además, un dispositivo para la autentificación de participantes en un servicio de telefonía que comprende un primer terminal de un participante conectado mediante un canal de transmisión a al menos un segundo terminal, una unidad de registro configurada para registrar las señales de voz transmitidas a través del canal de transmisión en el primer terminal y una unidad de procesamiento estando configurada la unidad de procesamiento para generar un primer conjunto de datos de verificación a partir de las señales de voz registradas. Además, el primer terminal está acoplado a un canal de seguridad a través del cual se trasmite un primer conjunto de datos de verificación y/o al menos un segundo conjunto de datos de verificación que se genera en el segundo terminal a partir de la señal de voz registrada en el mismo, en el que el conjunto de datos de verificación enviados a través del canal de seguridad se comprara con el otro.

En una realización preferida de la presente invención, el primer conjunto de datos de verificación se envía desde el primer terminal a través del canal de seguridad y el segundo terminal está acoplado al canal de seguridad para recibir el primer conjunto de datos de verificación y comparar el primer conjunto de datos de verificación recibido a través del canal de seguridad con el segundo conjunto de datos de verificación. Esto permite verificar en el segundo terminal la autenticidad del participante en el dispositivo según la invención.

Además, en otra realización preferida de la presente invención, el segundo conjunto de datos de verificación puede enviarse a través del canal de seguridad desde el segundo terminal y el primer terminal recibe el segundo conjunto de datos de verificación a través del canal de seguridad, en el que el dispositivo comprende, además, una unidad de comparación acoplada al primer terminal para comparar el segundo conjunto de datos de verificación recibido a través del canal de seguridad con el primer conjunto de datos de verificación. Esto permite realizar una verificación bidireccional de la autenticidad tanto del participante del dispositivo como del participante del segundo terminal. Mediante el intercambio de otros conjuntos de datos de verificación de otros terminales acoplados entre sí a través del canal de transmisión, en particular el participante del dispositivo puede realizar una verificación mutua de la autenticidad del resto de participantes y, por tanto, de la integridad del canal de transmisión.

En otra realización preferida de la presente invención, el dispositivo de registro registra un segmento de señal de voz que se divide en un conjunto de ventanas.

Preferentemente, el conjunto de datos de verificación generado a partir de las señales de voz comprende al menos una de las ventanas y/o las características de voz calculadas a partir de al menos una de las ventanas.

En una realización particularmente ventajosa de la presente invención, la unidad de comparación está configurada, además, para definir una selección determinística aleatoria del conjunto de ventanas y para definir una distancia entre los datos vinculados a las ventanas seleccionadas a partir del segundo conjunto de datos de verificación recibido a través del canal de seguridad y el primer conjunto de datos de verificación.

En otra realización preferida de la presente invención, el dispositivo comprende, además, un dispositivo de lectura acopado al terminal del participante que está configurado para leer las características de seguridad de una memoria de datos personalizada para el participante, estando conectado el dispositivo de lectura a un servidor y configurado para enviar las características de seguridad leídas al servidor a fin de determinar la identidad del participante.

Preferentemente, la memoria de datos está dispuesta en un documento de identidad y el dispositivo de lectura está configurado para leer los datos de dicha tarjeta de identidad.

Otras ventajas y características del procedimiento, sistema y dispositivo para la autentificación de participantes en un

servicio de telefonía según la invención se deducen de la siguiente descripción, en la que la invención se explica en mayor detalle mediante ejemplos de realización, así como en referencia a los dibujos adjuntos. Muestran:

La figura 1, un sistema de comunicación para la autentificación de participantes según una realización de la presente invención,
La figura 2, un diagrama de flujo temporal que representa un procedimiento según una realización de la presente invención con dos terminales,
La figura 3, un diagrama de flujo temporal que representa un procedimiento según otra realización de la presente invención con dos terminales,
La figura 4, un diagrama de flujo que representa un procedimiento de autentificación según una realización de la presente invención,
La figura 5, un diagrama de flujo temporal que ilustra la señal de voz registrada en cada caso, así como una división en ventanas y una comparación de las ventanas según una realización de la presente invención,
La figura 6A y la figura 6B, dos realizaciones de un conjunto de datos de verificación según realizaciones de la presente invención,
La figura 7, un diagrama de flujo que representa ciertos aspectos de una comparación de características según una realización de la presente invención, y
La figura 8, un diagrama de clases de una realización técnica mediante software según una realización de la presente invención.

La figura 1 muestra un sistema de comunicación para la autentificación de participantes en un servicio de telefonía según una realización de la presente invención. El sistema 1 comprende un canal de transmisión 3, que conecta un terminal 5 de un participante con una estación remota 7. El canal de transmisión 3 puede ser un canal telefónico y, en particular, un canal telefónico clásico basado en una centralita, analógico o digital, o permitir la telecomunicación basada en paquetes entre el terminal 5 y la estación remota 7. El participante 9 del terminal 5 puede comunicarse a través del terminal 5 con el participante 9' de la estación remota 7. El participante 9 puede ser, por ejemplo, un cliente de banca en línea y el participante 9' puede ser un empleado o representante de servicio al cliente de dicha banca en línea. Como se muestra en la figura 1, el terminal 5 puede ser un teléfono móvil, un teléfono inteligente, un pad, una tableta u otro dispositivo de comunicación móvil configurado para la transmisión de voz. El terminal 5 puede ser, además, un teléfono fijo, inalámbrico o por cable o un sistema telefónico a través del cual pueda comunicarse el participante 9. Cualquier terminal puede conectarse al canal de transmisión 3 en la estación remota 7, por ejemplo, un dispositivo telefónico por cable habitual o un sistema telefónico que el participante 9 pueda utilizar, por ejemplo, mediante unos auriculares. El terminal de la estación remota 7 también puede ser un sistema informático configurado para la comunicación, por ejemplo, un servidor de comunicación que se comunica con un teléfono del participante 9'. Resulta evidente que el terminal de la estación remota 7 puede ser también uno de los terminales 5. En particular, la presente invención no se limita específicamente a un terminal ni a una configuración específica de un terminal. Por el contrario, también son concebibles otras configuraciones y combinaciones posteriores y la invención no está limitada en este sentido.

El sistema 1 presenta, además, un canal de seguridad 11 que conecta el terminal 5 del participante 9 a la estación remota 7, además de al canal de transmisión 7, y que está configurado preferentemente para la transmisión de datos, por ejemplo, de datos basados en paquetes. Un canal de comunicación adicional 13 conecta, además, un lector 15 a un servidor 17. Para ello, el lector 15 puede estar configurado para leer una memoria de datos de un documento de identidad del participante 9 a fin de determinar la identidad del participante 9 en comunicación con el servidor 17. Además, el dispositivo de lectura 15 puede leer otras características de autentificación adicionales del participante 9 desde el documento de identidad o desde otra memoria de datos personalizada, por ejemplo, desde una tarjeta de firma. No obstante, el dispositivo de lectura 15 también puede estar diseñado como una interfaz integrada en un teléfono inteligente u otro dispositivo móvil, por ejemplo, y por lo tanto, como un componente del terminal 5. Asimismo, el dispositivo de lectura 15 puede estar diseñado como una interfaz NFC (Near Field Communication), por ejemplo, en un teléfono inteligente u otro dispositivo móvil, capaz de comunicarse directamente con un chip RFID del documento de identificación.

Como se muestra en la figura, los canales 3, 11 y 13 pueden configurarse de forma independiente entre sí y proporcionar una conexión, por ejemplo, mediante redes de comunicación independientes. No obstante, los canales 11 y 13 también pueden realizarse de forma conjunta mediante una infraestructura de seguridad. Por consiguiente, el canal de transmisión 3 también puede configurarse en la misma red de la infraestructura de seguridad.

La autenticidad de los participantes 9, 9', en particular tras una verificación inicial de la identidad del participante 9 mediante el servidor 17, se verifica de forma continua durante la comunicación mediante el registro de las señales de voz tanto en el terminal 5 como en la estación remota 7, generándose a partir de las respectivas señales de voz

registradas un conjunto de datos de verificación al que se incorporan las características de la voz y otras propiedades mensurables de la señal de voz. En particular, esto permite garantizar que la identidad de los participantes 9, 9' no cambie. Para ello, pueden tenerse en cuenta los mismos segmentos de la señal de voz más largos a ambos lados de la comunicación gracias a la sincronización realizada. En dichos segmentos se incluyen intervalos de tiempo o ventanas más cortos, que forman la unidad de evaluación más pequeña. Los conjuntos de datos de verificación recibidos se emiten al lado opuesto a través del canal de seguridad 11. En el lado opuesto, se comparan con el resultado del análisis obtenido de la llamada telefónica registrada en ese punto en las mismas condiciones marco.

El participante 9 puede configurar el terminal 5 para que el conjunto de datos de verificación local se envíe a la estación remota 7 a través del canal de seguridad 11. Cuando la estación remota 7 recibe el conjunto de datos de verificación desde el dispositivo terminal 5, puede comparar el conjunto de datos de verificación recibido con el conjunto de datos de verificación local. Para ello, la estación remota 7 puede acoplarse a una unidad de comparación 19 que recibe los dos conjuntos de datos de verificación y los comprueba para detectar coincidencias y discrepancias. Además, la unidad de comparación 19 puede estar configurada únicamente para recibir el conjunto de datos de verificación del terminal 5 a través del canal de seguridad 11 y reenviarlo al terminal o estación remota 7 del participante 9', de forma que el terminal o estación remota 7 puede comprobar si el conjunto de datos de verificación con el conjunto de datos de verificación recibido por la unidad de comparación 19 para detectar coincidencias y discrepancias. Si los dos conjuntos de datos de verificación coinciden, la estación remota 7 puede determinar que el participante 9 se considera auténtico e indicárselo al participante 9'. Esto permite, por ejemplo, que un proveedor de servicios, por ejemplo, de banca en línea pueda garantizar que se produzca una comunicación con el participante 9 inicialmente identificado de forma continua a través del canal de transmisión 3.

De forma alternativa o adicional, el conjunto de datos de verificación generado en la estación remota 7 también puede transmitirse al dispositivo terminal 5 del participante 9 a través del canal de seguridad 11. Para comparar el conjunto de datos de verificación transmitidos con el conjunto de datos de verificación local, el terminal 5 puede comprender una unidad de comparación integrada que puede instalarse y ejecutarse como un módulo de software en el terminal 5, acoplarse al terminal 5 como un módulo de hardware y/o acoplarse al terminal 5 como una combinación de hardware y software para comparar los conjuntos de datos de verificación. El terminal 5 puede utilizar la unidad de comparación para comparar el conjunto de datos de verificación recibido a través del canal de seguridad 11 con el conjunto de datos de verificación local y, a partir de esta comparación, determinar si el participante 9' de la estación remota 7 y, por lo tanto, el canal de transmisión 3, es auténtico. Mediante esta comparación, el participante 9 puede determinar si la telecomunicación se sigue produciendo con el mismo participante 9' y si el canal de transmisión 3 no ha sido puesto en peligro y redirigido por un tercero o por un posible atacante, por ejemplo, para obtener otras características de autentificación del participante 9.

Es posible establecer un canal cifrado mediante una conexión de red con ayuda de una infraestructura de seguridad, que puede proporcionarse para su uso con un documento de identidad electrónico. Así, en el sistema a modo de ejemplo 1 que se muestra en la figura 1, un cliente puede hacer una llamada telefónica a un banco. El cliente puede autentificarse previamente en el banco, por ejemplo, con un documento de identidad electrónico (nPA) mediante una infraestructura de seguridad nPA. Además, entre el cliente y el banco se establece una conexión segura a través del canal de seguridad 11, por ejemplo, mediante el servicio web de la infraestructura de seguridad de nPA. Ambos terminales 5, 7 pueden ejecutar un componente de software que analiza la señal de voz recibida durante la llamada en curso y calcula las propiedades características de un segmento limitado en el tiempo y conocido por ambas partes de la señal telefónica. Este componente de software puede implementar el procedimiento según la invención según una realización de la presente invención. Los componentes de software sincronizan los relojes a través de la conexión segura para poder determinar segmentos de señal temporalmente idénticos. Los parámetros temporales exactos de los segmentos de señal se modifican de forma determinística aleatoria, lo que aumenta la seguridad del procedimiento frente a posibles ataques. El componente del cliente puede calcular las propiedades características de su propia señal de voz, mientras que el componente del banco puede calcular las propiedades características de la señal de voz recibida por el cliente. Los componentes pueden intercambiar los resultados calculados del análisis a través de la conexión segura 11 y compararlos con los resultados de los análisis locales.

En particular en sistemas que permiten la comunicación centralizada con una pluralidad de clientes, por ejemplo, en el caso de clientes que se comunican con un sistema de comunicación o un centro de llamadas de un banco u otro proveedor de servicios, es necesario asignar una persona que llama o cliente a un participante en el lado del proveedor de servicios, por ejemplo, un agente del centro de llamadas. De este modo, las sesiones web mediante las cuales se realiza la autentificación, por ejemplo, con el documento de identidad electrónico a través de la infraestructura de seguridad nPA del banco, pueden asignarse directamente y con facilidad. La asignación de qué llamada telefónica entrante pertenece a las respectivas sesiones web puede realizarse, por ejemplo, transmitiendo

un identificador seguro a través del canal de transmisión 3 y/o del canal de seguridad 11, lo que puede conseguirse, por ejemplo, con una marca de agua de audio o una modulación del canal de voz o el canal de transmisión 3. De este modo, el sistema 1 permite tanto la autentificación unilateral del resto de participantes del servicio de telefonía como la autentificación multilateral por parte de cada uno de los participantes 9, 9', de modo que cada participante 9, 9' puede decidir de forma autónoma e independiente del resto si la telecomunicación sigue siendo fiable o si existe un cierto riesgo de que la identidad de los participantes en la comunicación y/o la integridad del canal de transmisión 3 pueda haber cambiado debido a un ataque contra el sistema 1, por ejemplo, un intento de suplantación de identidad, mediante la introducción de pulsos de interferencia selectivos o mediante otros ataques.

Por lo tanto, el sistema de comunicación 1 permite aumentar significativamente la seguridad de un servicio de telefonía y, en particular, una verificación de la integridad y autenticidad de un canal de comunicación, por ejemplo, del canal de transmisión 3, pudiendo la autentificación de los participantes 9, 9' según la invención basarse en cualquier sistema de comunicación existente, por ejemplo, un sistema de comunicación analógico, digital o basado en paquetes, por lo que no es necesario efectuar ningún cambio en dicho sistema de comunicación existente. Por el contrario, la autentificación según la invención puede integrarse fácilmente en cada terminal 5, 7 mediante módulos de software adecuados, por ejemplo, en forma de aplicaciones u otros programas, de modo que tanto el registro de las señales de voz como la comparación de los conjuntos de datos de verificación pueden ejecutarse mediante los recursos informáticos del terminal 5, 7. No obstante, también resulta evidente que, de forma alternativa o adicional, pueden preverse módulos de hardware o componentes de hardware, por ejemplo, en la estación remota 7, a fin de permitir un registro y procesamiento más eficiente de las señales de voz.

Además, el sistema 1 no se limita únicamente a dos participantes 9, 9'. Por el contrario, el sistema 1 puede conectar entre sí a un número cualquiera de participantes adicionales a través del canal de transmisión 3, por ejemplo, en una conferencia telefónica. En este caso, tanto el canal de transmisión 3 como el canal de seguridad 11 entre la estación remota 7 y cada uno de los terminales del resto de participantes pueden tener forma de estrella. No obstante, la presenta invención no se limita a una topología específica de los canales 3, 11, sino que puede comprender otras topologías, por ejemplo, topologías en anillo, o utilizar una o varias redes generales para la transmisión tanto de datos como de voz para conectar a los participantes entre sí. En una configuración con una pluralidad de participantes, el conjunto de datos de verificación correspondiente desde cada terminal puede enviarse a través del canal de seguridad 11 al resto de participantes, en el que se comparará con el conjunto de datos de verificación local. Alternativamente, puede enviarse únicamente el conjunto de datos de verificación del participante 9' desde la estación remota 7 al resto de participantes, o puede preverse que solo una cantidad parcial de participantes intercambien sus correspondientes conjuntos de datos de verificación a través del canal de seguridad 11 a fin de demostrar su autenticidad al resto de participantes en la comunicación.

La presente invención puede utilizarse en una pluralidad de sistemas de comunicación y procedimientos de transmisión, por ejemplo, con telefonía analógica, RDSI, ADSL o VDSL, pero también con procedimientos inalámbricos como GSM, GSM con HSCSD o GPRS, EDGE o HSDPA. Además, la presente invención puede integrarse en un teléfono inteligente u otro dispositivo móvil adecuado, por ejemplo, una tableta o pad, o puede comprender un terminal clásico, por ejemplo, un teléfono, y una unidad de procesamiento, por ejemplo, un ordenador de escritorio u ordenador portátil. El enlace de comunicación entre las dos unidades de procesamiento puede ser una conexión Ethernet o cualquier otra conexión segura basada en paquetes. No obstante, los sistemas de procesamiento también pueden utilizar la red que conecta los terminales para establecer el canal de seguridad. Por lo tanto, la presente invención no está limitada, en particular, por un terminal específico 5, 7 ni por una infraestructura de comunicación específica.

La figura 2 muestra un diagrama de flujo temporal que representa el desarrollo de un procedimiento según una realización de la presente invención con dos participantes. Se muestra un cliente 21 y un servidor 23, que pueden corresponderse, por ejemplo, con el terminal 5 y la estación remota 7 de la figura 1. Durante el inicio, se sincronizan el cliente 21 y el servidor 23 sincronizando los relojes tanto del cliente 21 como del servidor 23 utilizando, por ejemplo, el protocolo Network Time Protocol (NTP). A continuación, la demora de transmisión del canal de transmisión entre el cliente 21 y el servidor 23 se determina gracias a la definición 25 del tiempo de ida y vuelta (RTT) mediante el envío de un mensaje, por ejemplo, un ping, desde el cliente 21 al servidor 23 y de vuelta al cliente 21. A partir de esto, se obtiene inmediatamente según $RTT = 2 \cdot STT$ el tiempo en un solo sentido (STT), que especifica un valor para la demora de transmisión entre el cliente 21 y el servidor 23. No obstante, dado que esta definición 25 puede variar en función de la capacidad del canal de transmisión, la definición 25 puede repetirse varias veces 27 a fin de poder estimar con mayor precisión una demora de transmisión media. La definición del valor de demora STT permite garantizar que la señal de voz registrada en el cliente 21 se registre también en el servidor 23 esencialmente en el mismo momento. Asimismo, pueden preverse procedimientos o componentes para compensar los errores que puedan producirse debido a una mala sincronización temporal de los participantes y, por

lo tanto, del cliente 21 y el servidor 23.

Después de las etapas 25, 27, el cliente 21 envía un valor de inicio para inicializar un generador de números pseudoaleatorios que está presente tanto en el cliente 21 como en el servidor 23. A continuación, el cliente 21 define un momento de inicio t_0 para la grabación de la señal de voz y envía este valor al servidor 23. No obstante, resulta evidente que el servidor 23 también puede definir el momento de inicio t_0 y comunicarse, por tanto, con el cliente 21.

El Cliente 21 empieza a registrar la señal de voz 29 en el momento t_0. En primer lugar, el cliente 21 registra un segmento de la señal de voz 31, que puede tener una longitud de, por ejemplo, Δt_m= 1,6 s. La señal de voz disponible 31' se registra también en el servidor 23, pero demorada por el valor de demora calculado STT. También en este caso, el segmento de la señal de voz 31' puede presentar una longitud de, por ejemplo, Δt_m= 1,6 s. Tanto en el cliente 21 como en el servidor 23, el segmento de la señal de voz 31, 33 registrado correspondiente se divide en un conjunto de ventanas, de las cuales solo dos se muestran en la figura 2. Tanto el cliente 21 como el servidor 23 definen mediante el generador local de números pseudoaleatorios un conjunto de números pseudoaleatorios que se utilizan para identificar las ventanas de los segmentos de la señal de voz 31, 31' que van a utilizarse para el análisis y procesamiento posteriores. Como se explicó anteriormente, estos números pseudoaleatorios son siempre los mismos para el cliente 21 y el servidor 23, ya que los generadores de números pseudoaleatorios se inicializaron con el mismo valor de inicio. Las características se extraen a partir de las ventanas definidas de forma determinística aleatoria por ambos lados y se utilizan para generar los conjuntos de datos de verificación 35 o 35'. De forma paralela al análisis de las características y la generación de los conjuntos de datos de verificación 35, 35', tanto el cliente 21 como el servidor 23 pueden seguir registrando otros segmentos de la señal de voz 37, 37', por ejemplo, con una demora de Δt_{delay}= 0,1 s, lo que garantiza la verificación continua de la autenticidad del canal de transmisión y de los participantes.

En la realización que se muestra en la figura 2, el servidor 23 envía el conjunto de datos de verificación 35' al cliente 21, de modo que el cliente 21 puede determinar, mediante la comparación 39 de los conjuntos de datos de verificación 35, 35', si el servidor 23 y, por lo tanto, el canal de transmisión, al menos en parte, es auténtico o si existe un posible riesgo para la seguridad.

La figura 3 muestra un diagrama de flujo temporal de una autentificación según otra realización de la presente invención, en la que el procedimiento mostrado permite una verificación mutua de la autenticidad de los participantes implicados. El procedimiento también puede realizarse mediante los componentes que aparecen en la fig. 2, por lo que los componentes que coinciden aparecen marcados con las mismas referencias. La figura 3 muestra, de forma similar a la figura 2, un cliente 21 y un servidor 23 que se comunican entre sí a través de un canal de transmisión y utilizan un canal de seguridad para verificar la autenticidad del interlocutor. En la realización según la figura 3, la autentificación y la verificación de identidad 41 se realizan, en primer lugar, mediante una infraestructura de seguridad. A continuación, el cliente 21 y el servidor 23 negocian las respectivas funciones y se sincronizan 43 los relojes utilizando, por ejemplo, el protocolo NTP. En la realización que se muestra en esta figura y en base a la distribución de las funciones, el servidor 23 envía un valor de inicio para inicializar los generadores de números pseudoaleatorios de ambos lados y fija un momento de inicio del registro de las señales de voz 45. Dependiendo del momento de inicio, tanto el cliente 21 como el servidor 23 registran las señales de voz correspondientes de forma continua y sincrónica en los segmentos de señal de voz 31, 37 en el lado del cliente y 31', 37' en el lado del servidor. Basándose en el segmento de la señal de voz 31, el cliente 21 genera un conjunto de datos de verificación 35 y el servidor 23 genera, a su vez, el conjunto de datos de verificación 35' a partir del segmento de la señal de voz 31'. No obstante, a diferencia de la realización que se muestra en la figura 2, tanto el cliente 21 como el servidor 23 envían el conjunto de datos de verificación local 35, 35' a través del canal de seguridad al respectivo lado opuesto, de forma que, mediante una comparación 47 de los conjuntos de datos de verificación 35, 35', puede determinarse, en cada lado y de forma independiente, si la integridad del canal de transmisión está garantizada y si la autenticidad del otro participante en la comunicación es constante.

La figura 4 muestra un diagrama de flujo de un procedimiento de autentificación según una realización de la presente invención, como puede realizarse en un terminal, por ejemplo, en el servidor 23 que aparece en las figuras 2 y 3. Los distintos bloques del diagrama representan componentes funcionales que pueden ejecutarse, desde un registro hasta una decisión de clasificación. El procedimiento comienza con un registro 49 de un segmento de la señal de voz de las señales de voz locales. La señal registrada se somete a un procesamiento previo adecuado 51, por ejemplo, normalizándola y filtrándola adecuadamente. El segmento de la señal de voz sometido a un procesamiento previo, a continuación, en la etapa 53, puede dividirse en un conjunto de ventanas que pueden cubrir preferentemente todo el segmento de señal de voz. Las distintas ventanas pueden tener una longitud de Δt_F y pueden ser directamente adyacentes (O_F= 0) o solaparse hasta cierto punto, por ejemplo a la mitad (O_F= 0,5), dependiendo de un factor de superposición O_F 0 <O_F< 1.

Cada ventana puede someterse a un procesamiento adicional, por ejemplo, aplicando un filtro de Hamming. Basándose en estos datos, a continuación se extrae una pluralidad de características 55 que se almacenan localmente en un conjunto de datos de verificación y se envían al resto de participantes a través de un canal de seguridad 57, recibiéndose también conjuntos de datos de verificación del resto de participantes en la etapa 57. Como se muestra en la fig. 4, los conjuntos de datos de verificación pueden intercambiarse como vectores de características, pudiendo comprender un conjunto de datos de verificación uno o más vectores de características para cada ventana.

La señal de cada ventana sometida a un procesamiento previo puede transformarse en un espacio de frecuencias mediante una transformada de Fourier o una transformada rápida de Fourier (FFT). A partir de los coeficientes de Fourier y de la señal temporal original, pueden calcularse otras características y vectores de características, por ejemplo, características analíticas y estadísticas, como compacidad, flujo espectral, variabilidad espectral, coeficientes cepstrales en las frecuencias de Mel o momentos de una distribución estadística, entre otros.

El algoritmo de extracción para el cálculo de la compacidad puede calcular, por ejemplo, una medida de la componente de ruido de un segmento de señal. Así, el grado de compacidad de un segmento de señal dado puede determinarse a partir de un espectro de magnitud $|X(k)| = \{k_0, k_1, ..., k_N\}$ previamente calculado mediante una FFT. En este caso, N es el número de frecuencias. k_i es la magnitud de la frecuencia f_i con $k_i \in R$. En este caso, puede ejecutarse un bucle con índice de ejecución i y $0 \leq i \leq N$ mediante $|X(k)|$. Para valores de i, en los que la magnitud de las tres potencias adyacentes k_{i-1}, k_i, k_{i+1} es superior a 0, la compacidad C de todo el segmento de señal se puede calcular, por ejemplo, como:

$$C = \sum_{0}^{N} \left| \frac{20 \cdot \log k_i - 20 \cdot \left(\log k_{i-1} + \log k_i + \log k_{i+1} \right)}{3} \right| .$$

El flujo espectral puede ser una medida de la variación espectral de una señal. Para ello, se pueden sumar los cuadrados de las diferencias de los espectros de magnitud de dos ventanas sucesivas. El espectro de magnitud $|X(k)|$ puede ser el espectro de la ventana actual y el espectro de magnitud $|X'(k)|$ puede ser el espectro de la ventana anterior. En ambos casos, N puede especificar el número de frecuencias en el espectro de magnitud. A continuación, el flujo espectral F_X del algoritmo se puede calcular de la siguiente manera:

$$F_X = \sum_{1}^{N} \left(k_i - k'_i \right)^2 .$$

Además, la variabilidad espectral V_X de un segmento de señal, que puede representar una desviación estándar estadística σ del espectro de magnitud $|X(k)|$, se puede calcular de la siguiente manera:

$$V_X = \sigma\left(|X(k)| \right) = \sqrt{\frac{1}{N} \cdot \sum_{i=0}^{N} \left(k_i - \overline{X} \right)^2} \text{ , en el que } \overline{X} = \frac{1}{N} \cdot \sum_{i=0}^{n} k_i .$$

Para calcular otras características, puede utilizarse el método del algoritmo del Método de los momentos o "Method of Moments", que puede extraer características a partir de segmentos de señales dados utilizando medios estadísticos. Para ello, también se tiene en cuenta el espectro de magnitud $|X(k)|$ de un segmento de señal. Se calculan los momentos de una distribución estadística que son característicos de dicha distribución. Para ello, puede formarse en primer lugar, por ejemplo, la suma S_X del espectro de magnitud $|X(k)|$, que puede registrarse como área bajo una curva que se puede interpretar como el espectro de magnitud, en el que

$$S_X = \sum_{1}^{N} k_i$$

A continuación, pueden normalizarse todos los valores de $|X(k)|$ con S_X a $|X(k)_{Norm}|$ y, a continuación, puede calcularse $M_{m,temp}$ como:

$$M_{m,temp(i)} = \sum_{1}^{N} i \cdot k_{Norm(i)} .$$

Por último, puede determinarse un valor de salida M_m con $M_m = \{M_1, M_2, ..., M_5 | M_{mi} \in R\}$, en el que:

$$M_1 = S_X ,$$

$$M_2 = M_{m,temp(2)} ;$$

$$M_3 = M_{m,temp(3)} - M_{m,temp(2)}^{\ 2} ,$$

$$M_4 = 2 \cdot M_{m,temp(2)}^{\ 3} - 3 \cdot \left(M_{m,temp(2)} \cdot M_{m,temp(3)} \cdot M_{m,temp(4)} \right)$$

y

$$M_5 = -3 \cdot M_{m,temp(2)}^{\ 4} + 3 \cdot \left(M_{m,temp(2)} \cdot M_{m,temp(2)} \cdot M_{m,temp(3)} \right) - 4 \cdot \left(M_{m,temp(2)} \cdot M_{m,temp(4)} \cdot M_{m,temp(5)} \right) .$$

No obstante, resulta evidente que la presente invención no se limita a los procedimientos de extracción de características descritos, sino que puede extraerse y aplicarse cualquier característica adecuada que pueda utilizarse para caracterizar señales de voz y de audio a fin de poder sacar conclusiones sobre el contenido y las propiedades de la información de voz.

Además, los distintos filtros de extracción de características pueden organizarse en un banco de filtros y puede procesarse la señal de entrada. Al calcular las características individuales, pueden utilizarse valores determinísticos aleatorios para los distintos parámetros del procedimiento de extracción de características, que pueden calcularse a partir de un valor de inicio intercambiado inicialmente entre los participantes. Dado que el número de parámetros para calcular las características de voz correspondientes se conoce de antemano y es el mismo en cada terminal, gracias a la definición de números pseudoaleatorios, siempre se determinan los mismos valores, que son (seudo) aleatorios, para la parametrización.

Tras el intercambio de vectores de características en la etapa 57, el servidor 23 compara los vectores de características recibidos por el cliente 21 en la etapa 59. Aunque la etapa 59 se representa únicamente para la comparación de vectores de características de dos participantes, la presente invención no se limita a solo dos participantes. Por el contrario, el servidor 23 puede comparar, secuencialmente o en paralelo, los vectores de características de varios participantes con sus propios vectores de característica y confirmar o no la autenticidad del participante correspondiente a partir del resultado.

El servidor 23 puede calcular un resultado global 61 a partir de los resultados de la comparación 59 y basándose en dicho resultado, tomar una decisión 63 sobre la integridad del canal de transmisión de la comunicación y sobre la autenticidad del resto de participantes, ya sea mediante una simple comparación de valores de umbral o utilizando clasificadores u otros procedimientos apropiados.

La figura 5 muestra una sección del diagrama de flujo temporal de la figura 3 que muestra la división de un segmento de señal de voz registrado 31, 31' en distintas ventanas F_1, ..., F_{15}. Además, la figura 5 ilustra una comparación de las distintas ventanas según una realización de la presente invención. Para dificultar posibles ataques contra el procedimiento según la invención, los índices de tiempo de las ventanas en cuestión se determinan de forma determinística aleatoria en los segmentos de la señal de voz 31, 31' correspondientes. Tras el registro de los segmentos de la señal de voz 31, 31' y de la distribución en distintas ventanas F_1, ..., F_{15}, pueden determinarse de forma determinística aleatoria índices de tiempo de las ventanas en cuestión. Para cada intervalo de tiempo, también en este caso se utiliza un generador de números pseudoaleatorios para calcular nuevos números pseudoaleatorios en cada terminal a partir de un valor de inicio distribuido inicialmente.

Por una parte, pueden analizarse las ventanas de un primer grupo que se corresponden en el tiempo, por ejemplo, las ventanas F_2, F_8 y F_{10} de los segmentos de señal 31 y 31'. Estos presentan el mismo índice de tiempo para ambos participantes y sus vectores de características tienen la misma longitud debido a la parametrización. Debido a la correspondencia temporal de estas ventanas, una comparación debería demostrar una equivalencia elevada. A partir de esto, se obtiene inmediatamente un criterio para la autenticidad del otro participante. Por otra parte, preferentemente se compararán también las ventanas que no se corresponden en el tiempo. Estas ventanas difieren en su índice de tiempo, pero por lo demás, presentan los mismos parámetros que las ventanas del primer grupo, que

se verificaron en busca de equivalencias. Por ejemplo, en el segundo grupo se comparan las ventanas F_1 y F_{10}, F_2 y F_5, así como F_8 y F_{14}. Si la transmisión no presenta interferencias, estos segmentos de señal deberían identificarse de forma fiable como no idénticos.

La comparación de ventanas no idénticas puede aumentar significativamente la calidad de asignación y, por lo tanto, la seguridad del sistema. Gracias a esto se logra, en particular, una mayor tolerancia a los fallos en comparación con una entrada de señal continua. Las comparaciones, tanto de ventanas idénticas como distintas, en busca de equivalencias y no equivalencias, se incorporan preferentemente al resultado global que se utilizará para evaluar la integridad del canal de comunicación y la autenticidad de los participantes. Por ejemplo, solo deberá entenderse que el canal de transmisión es auténtico si las ventanas de tiempo síncronas son idénticas y las ventanas de tiempo distintas no son idénticas. En todos los demás casos, puede entenderse que la integridad del canal de transmisión se ha visto perturbada y, por lo tanto, no puede darse por hecho que la comunicación sea segura. Además, los resultados de las comparaciones en busca de equivalencia y no equivalencia pueden ponderarse de forma diferente. Por ejemplo, el resultado de la comparación en busca de equivalencia puede recibir un factor de ponderación más elevado, ya que la equivalencia de ventanas temporalmente idénticas puede ser un criterio más importante o significativo para la autenticidad del canal. Por otra parte, la no coincidencia de las ventanas de las ventanas temporalmente distintas puede ser importante para la verificación de errores, aunque en general, es secundaria.

Mediante la determinación determinística aleatoria de las ventanas a comparar F_i para cada segmento de la señal de voz 31, 31', según realizaciones de la presente invención, pueden utilizarse distintos conjuntos de datos de verificación para la transmisión, como se muestra en las figuras 6A y 6B. Por ejemplo, en el conjunto de datos de verificación a transmitir, solo pueden integrarse y transmitirse al resto de participantes a través del canal de seguridad ventanas seleccionadas. Para ello, puede generarse, como se muestra en las figs. 6A y 6B, un vector índice 65 que comprenda varias tuplas $(i, j)_k$ de índices. La primera entrada i de la tupla k $(i, j)_k$ da el índice de una ventana F_i, que debe verificarse en busca de coincidencia, y la segunda entrada j indiza otra ventana F_j, que debe comprobarse con la primera ventana F_i en busca de no coincidencia. Por ejemplo, una tupla (2, 5) define una comparación de las ventanas F_2 en busca de coincidencia y una comparación de la ventana local F_2 con la ventana remota F_5 en busca de no coincidencia.

En estos casos, siempre se puede acceder a los datos locales A_L67 a través de su índice, ya que pueden estar disponibles en su totalidad. No obstante, los datos característicos remotos y los resultados del análisis A_R69, 69' pueden diferir en las realizaciones que se muestran en la fig. 6A y 6B. Si, como se muestra en la fig. 6B, todos los datos característicos 69' están disponibles también en este caso, se accederá a los mismos a través del índice. Por el contrario, si, como se muestra en la fig. 6A, solo se ha transmitido una selección de los datos característicos remotos 69, puede accederse a los datos característicos a través del índice k de la tupla $(i, j)_k$ en el vector de índice 65. En este caso, se aplica lo siguiente: para la comparación de coincidencia de las ventanas F_i, pueden seleccionarse los datos característicos remotos con índice $2 \cdot k$ y para la comparación de no coincidencia con la ventana F_j, pueden utilizarse los datos característicos remotos con índice $2 \cdot k + 1$ del vector característico 69.

La figura 7 muestra un diagrama de flujo para la comparación del conjunto de datos de verificación recibido con el conjunto de datos de verificación local según una realización de la presente invención, como se muestra, por ejemplo, en las etapas 39 de la figura 2, 47 de la figura 3 y 59 a 63 de la figura 4. Como se explicó anteriormente, para cada ventana pueden calcularse distintas características e integrarlas en el conjunto de datos de verificación. Esta pluralidad de características puede compararse durante la comparación del conjunto de datos de verificación local con el conjunto de datos de verificación recibido. En la figura 7, se representa a modo de ejemplo la comparación de las características flujo espectral 71a, compacidad 71b y momentos estadísticos 71c. No obstante, resulta evidente que también pueden compararse otras características contenidas en los conjuntos de datos de verificación. Los conjuntos de datos se comparan entre sí 73a, 73b, 73c utilizando, por ejemplo, un procedimiento de distorsión temporal dinámica (DTW).

La comparación puede realizarse, por ejemplo, como una arquitectura paralela. Los valores de entrada 71a, 71b, 71c se agregan a los dos extractores de características D_i, cada uno de los cuales proporciona un resultado de la verificación de coincidencia y no coincidencia. De este modo, en la comparación 73a, 73b, 73c se generan dos valores de comparación c_{i1}, c_{i2}, que, a continuación, se comparan con un valor de umbral en la etapa 75, 75'. Como resultado, se calculan las variables cmp_{i1} y cmp_{i2}, cada una de los cuales indica si c_{i1} está por debajo del valor de umbral o si c_{i2} está por encima del valor umbral correspondiente. Por lo tanto, la variable cmp_{i1} contiene el valor 1 si c_{i1} está por debajo del valor de umbral y la verificación de los segmentos de señal coincidentes se ha realizado con éxito. Del mismo modo, cmp_{i2} contiene el valor de la comprobación de no coincidencia, en la que c_{i2} debe estar por encima del valor de umbral para que la verificación de los segmentos de señal no coincidentes se realice con éxito.

En las siguientes etapas, las variables de resultados cmp_{i1} y cmp_{i2} se combinan para formar un valor $comb_i$. La combinación de las variables de resultados cmp_{i1} y cmp_{i2} puede realizarse, por ejemplo, mediante una operación lógica AND. No obstante, preferentemente, las variables de resultados cmp_{i1} y cmp_{i2} de la verificación de coincidencia o de no coincidencia pueden ponderarse y solo entonces combinarse entre sí, por ejemplo, como una suma normalizada. A continuación, el valor $comb_i$ resultante puede ponderarse, a su vez, con un peso g_i, de modo que se obtiene el resultado $res_i = comb_i * b_i$ para una característica determinada en la etapa 77. Los resultados calculados de este modo de todos los clasificadores D_i 73a, 73b, 73c se suman en la etapa 78 y se verifica si se cumple una condición de éxito de la de votación por mayoría ponderada $w_{\Delta tm}$ con $\sum res_i > Valor\ de\ umbral_{Vote}$. El último componente 79 reúne todas las decisiones de votación por mayoría ponderada para el intervalo de análisis Δ_{tm}. Si el número de todas las decisiones $\sum w_{\Delta tm}$ supera un nivel de umbral, el canal de transmisión se clasificará como auténtico. La determinación de este valor de umbral, y por lo tanto, la determinación del número decisiones incorrectas por intervalo de análisis Δt_m que se considera aceptable para probar la autenticidad de una conexión depende preferentemente de una relación entre la robustez y la seguridad potencial del procedimiento.

La figura 8 representa un diagrama de clases que implementa componentes del sistema según la invención mediante software según una realización de la presente invención.

Al iniciar el sistema, se crea una instancia de clase MainGUI 81 y se accede al objeto correspondiente. La funcionalidad de este objeto consiste esencialmente en dos grupos, el acceso a los algoritmos de extracción y una concordancia de patrones. Para la extracción de características, se utiliza un objeto de clase DataModel 83 en el que se definen los parámetros de extracción. A continuación, este inicia una instancia de clase FeatureProcessorOpti 85 que controla el proceso de extracción y proporciona todas las funcionalidades necesarias para el mismo que no estén incluidas en los algoritmos de extracción. Por último, se accede de forma polifórmica a los algoritmos de extracción mediante FeatureProcessorOpti 85, en el que el prototipo de todos los algoritmos de extracción se define mediante FeatureExtractor 87.

Asimismo, el objeto MainGUI también puede instanciar los componentes que controlan la funcionalidad principal real del sistema: Las clases de los distintos participantes en la comunicación según sus funciones de cliente o servidor mediante las clases RTTServer 89 y RTTClient 91. Ambos utilizan la clase NetworkCommunicator 93 para comunicarse a través de la red. Por último, RTTClient instancia un objeto DecisionProcessor 95. Este encapsula todas las funciones relacionadas con la comparación de los datos de característica obtenidos. Éstos, a su vez, constan de dos grupos: Por un lado, el acceso a un clasificador, por ejemplo, FastDTW, y por otro, otros algoritmos de distorsión temporal dinámica para la concordancia de patrones. Dado que FastDTW no tiene por qué contener una clase superior unívoca ni una interfaz unívoca, TimeWarpInfo solo se muestra a modo de ejemplo. Por otro lado, el objeto DecisionProcessor contiene la funcionalidad para la evaluación de los resultados calculados mediante FastDTW a fin de decidir sobre la autenticidad de los participantes.

El diagrama de clases que se muestra en la figura 8 corresponde a una representación de un lenguaje de programación orientado a objetos y comprende componentes que son característicos para el lenguaje de programación Java en particular. No obstante, resulta evidente que la presente invención también puede implementarse mediante otros lenguajes de programación orientados a objetos, por ejemplo, Smalltalk, C++, C# o Pascal entre otros, y que no está limitada a las clases que se muestran en la figura 8. Por el contrario, una implementación concreta puede contener otras clases distintas y clases adicionales. También es posible una implementación mediante otros paradigmas de programación, por ejemplo, mediante un lenguaje de programación funcional o un lenguaje de programación lógico.

La presente invención puede descargarse como código fuente o como un programa ejecutable o como un componente de un programa ejecutable al dispositivo terminal correspondiente o leerse localmente a partir de una memoria de datos y compilarse e instalarse en consecuencia. El componente instalado correspondiente establece, en primer lugar, una conexión, por ejemplo, mediante el objeto instanciado NetworkCommunication 93, con el resto de componentes, mediante la cual pueden negociarse las distintas, por ejemplo, mediante los objetos MainGUI 81, y, a continuación, puede transferirse el procesamiento a la instancia de clase RTTServer 89 o RTTClient 91. No obstante, son concebibles otras realizaciones de la implementación técnica mediante software de esta realización de la presente invención y la presente invención no está limitada a una transmisión y realización en un determinado componente de procesamiento. Por el contrario, los distintos componentes y etapas de procesamiento ilustrados mediante el diagrama de clases pueden realizarse como componentes individuales de software o de hardware o como combinaciones de software y hardware. En particular, los algoritmos también pueden formarse exclusivamente en hardware.

Las características descritas en la descripción anterior, en las reivindicaciones y las figuras pueden ser relevantes

para la ejecución de la invención en sus realizaciones versiones, tanto individualmente como en cualquier combinación de las mismas.

Lista de referencias

1	Sistema
3	Canal de transmisión
5	Terminal
7	Estación remota
9,9'	Participante
11	Canal de seguridad
13	Canal de comunicación
15	Lector
17	Servidor
19	Unidad de comparación
21	Cliente
23	Servidor
25	Definición del tiempo de circulación
27	Repetición de la determinación
29	Registro de la señal de voz
31, 31'	Segmento de la señal de voz
35, 35'	Conjunto de datos de verificación
37, 37'	Segmento adicional de la señal de voz
39	Comparación de los conjuntos de datos de verificación
41	Autentificación / verificación de identidad
43	Reparto de funciones y sincronización de relojes
45	Transmisión de los valores iniciales
47	Comparación de los conjuntos de datos de verificación
49	Registro de las señales de voz
51	Procesamiento previo de la señal
53	Distribución del segmento de la señal de voz en ventanas
55	Extracción de las características
57	Intercambio de conjuntos de datos de verificación a través del canal de seguridad
59	Comparación de los conjuntos de datos de verificación
61	Cálculo del resultado global
63	Decisión sobre la autenticidad de los participantes
65	Vector índice
67	Conjunto de datos de verificación local
69, 69'	Conjunto de datos de verificación transmitido
71a, 71b, 71c	Datos de entrada de diversas características de voz
73a, 73b, 73c	Comparación de datos locales y remotos
75, 75'	Comprobación de identidad / no identidad
77	Cálculo y ponderación del resultado
78	Clasificación de los resultados
79	Autentificación del canal / de los participantes
82	Clase MainGUI
83	Clase DataModel
85	Clase FeatureProcessorOpti
87	Clase FeatureExtractor
89	Clase RTTServer
91	Clase RTTClient
93	Clase NetworkCommunication
95	Clase DecisionProcessor

REIVINDICACIONES

1. Procedimiento para la autentificación de participantes en un servicio de telefonía que comprende las etapas siguientes:

Establecimiento de un canal de transmisión entre un primer (23) y al menos un segundo (21) terminal de los participantes en el servicio de telefonía,

caracterizado por

el registro de las señales de voz que se transmiten a través del canal de transmisión en cada terminal (21, 23),
la generación de un conjunto de datos de verificación (35, 35') a partir de las señales de voz registradas en cada terminal (21, 23),
el envío de al menos un primer conjunto de datos de verificación (35') del primer terminal (23) desde el primer terminal (23) por un canal de seguridad y
la comparación (39) del el al menos un primer conjunto de datos de verificación (35') enviado a través del canal de seguridad con el al menos un segundo conjunto de datos de verificación (35) procedente del segundo terminal (21), en el que las señales de voz registradas en el primer terminal (23), a partir de las cuales se general al menos un primer conjunto de datos de verificación (35'), se corresponden esencialmente con las señales de voz registradas en el segundo terminal (21), a partir de las cuales se genera al menos un segundo conjunto de datos de verificación (35).

2. Procedimiento según la reivindicación 1, que comprende, además, la recepción del primer conjunto de datos de verificación (35') por el segundo terminal (21) a través del canal de seguridad, en el que se compara el primer conjunto de datos de verificación (35') recibido a través del canal de seguridad con el segundo conjunto de datos de verificación (35) mediante el segundo terminal (21).

3. Procedimiento según la reivindicación 1 o 2, que comprende, además, el envío del segundo conjunto de datos de verificación (35) a través del canal de seguridad desde el segundo terminal (21) , la recepción del segundo conjunto de datos de verificación (35) a través del canal de seguridad en el primer terminal (23) y la comparación del segundo conjunto de datos de verificación (35) recibido a través del canal de seguridad con el primer conjunto de datos de verificación (35') mediante el primer terminal (23).

4. Procedimiento según la reivindicación 1, que comprende, además, el envío del segundo conjunto de datos de verificación (35) desde el segundo terminal (21) a través del canal de seguridad, en el que se compara el primer conjunto de datos de verificación (35') enviado a través del canal de seguridad con el segundo conjunto de datos de verificación (35) enviado a través del canal de seguridad mediante una unidad de comparación que está acoplada al canal de seguridad y que recibe los dos conjuntos de datos de verificación enviados (35, 35').

5. Procedimiento según cualquiera de las reivindicaciones anteriores, que comprende, además, la fijación de un momento de inicio del registro de las señales de voz por uno de los terminales (21, 23), en el que las señales de voz se registran a partir del momento de inicio en los terminales (21, 23), esencialmente en el mismo momento o en un momento posterior, pospuesto por un valor de demora.

6. Procedimiento según cualquiera de las reivindicaciones anteriores, en el que el registro de las señales de voz comprende el registro de un segmento de la señal de voz (31, 37, 31', 37') que está dividido en un conjunto de ventanas, en el que cada conjunto de datos de verificación (35, 35') comprende al menos una de las ventanas.

7. Procedimiento según cualquiera de las reivindicaciones anteriores, en el que la generación del conjunto de datos de verificación (35, 35') comprende un cálculo de las características de voz de al menos una de las ventanas y la incorporación de las características de voz al conjunto de datos de verificación (35, 35') y la comparación (39) comprende una definición de una selección determinística aleatoria del conjunto de ventanas y una definición de una distancia entre los datos vinculados a las ventanas seleccionadas a partir del primer conjunto de datos de verificación (35') enviados a través del canal de seguridad y el segundo conjunto de datos de verificación (35).

8. Procedimiento según cualquiera de las reivindicaciones anteriores, que comprende, además, la determinación (41) de la identidad de un participante mediante la lectura de las características de seguridad a partir de una memoria de datos personalizada para el participante y la transmisión de las características de seguridad a un servidor, en el que la lectura comprende la lectura de los datos de un documento de identidad.

9. Soporte legible por ordenador con comandos almacenados en el mismo que, al ser ejecutado por un sistema informático, ordena al sistema informático que ejecute el procedimiento según cualquiera de las reivindicaciones anteriores.

10. Sistema para la autentificación de participantes en un servicio de telefonía que comprende lo siguiente:

un canal de transmisión (3), que conecta un primer terminal (5) y al menos un segundo terminal (7) de los participantes (9, 9') en el servicio de telefonía,

caracterizado por

un dispositivo de registro en cada terminal (5, 7), que está configurado el para registrar, en el terminal correspondiente (5, 7), las señales de voz transmitidas a través del canal de transmisión (3),
una unidad de procesamiento en cada terminal (5, 7), que está configurada para generar un conjunto de datos de verificación correspondiente a partir de las señales de voz registradas,
un canal de seguridad (11) acoplado al primer terminal (5), a través del cual se envía al menos un primer conjunto de datos de verificación del primer terminal (5), y
una unidad de comparación (19) configurada para comparar al menos un primer conjunto de datos de verificación enviado a través del canal de seguridad (11) con al menos un segundo conjunto de datos de verificación procedente del segundo terminal (7), en el que las señales de voz registradas en el primer terminal (5), a partir de las cuales se general al menos un primer conjunto de datos de verificación, se corresponden esencialmente con las señales de voz registradas en el segundo terminal (7), a partir de las cuales se genera al menos un segundo conjunto de datos de verificación.

11. Sistema según la reivindicación 10, en el que el segundo terminal (7) está acoplado al canal de seguridad (11) y a la unidad de comparación (19) para recibir el primer conjunto de datos de verificación y comparar el primer conjunto de datos de verificación recibido a través del canal de seguridad (11) con el segundo conjunto de datos de verificación y/o en el que el segundo terminal (7) está configurado, además, para enviar el segundo conjunto de datos de verificación a través del canal de seguridad (11), y el primer terminal (5) está configurado, además, para recibir el segundo conjunto de datos de verificación a través del canal de seguridad (11), en el que el sistema, además, comprende una unidad de comparación adicional acoplada al primer terminal (5) para comparar el segundo conjunto de datos de verificación recibido a través del canal de seguridad (11) con el primer conjunto de datos de verificación del primer terminal (7).

12. Sistema según la reivindicación 10 u 11, en el que uno de los terminales (5, 7) fija un momento de inicio para el registro y los dispositivos de registro registran las señales de voz a partir del momento de inicio en los terminales (5, 7), esencialmente en el mismo momento o en un momento posterior, pospuesto por un valor de demora, en el que el dispositivo de registro registra un segmento de señal de voz que se divide en un conjunto de ventanas y cada conjunto de datos de verificación comprende al menos una de las ventanas, en el que el conjunto de datos de verificación generado a partir de las señales de voz comprende las características de voz calculadas a partir de al menos una de las ventanas.

13. Sistema según cualquiera de las reivindicaciones 10 a 12, en el que el canal de transmisión (3) es un canal telefónico analógico o digital o proporciona una conexión basada en paquetes, preferentemente una conexión de telefonía por internet o una conexión de voz sobre IP.

14. Sistema según cualquiera de las reivindicaciones 10 a 13, en el que el sistema comprende, además, al menos un dispositivo de lectura (15) acopado a un terminal (5) de un participante (9) que está configurado para leer las características de seguridad de una memoria de datos personalizada para el participante (9), en el que el dispositivo de lectura (15) está conectado a un servidor (17) y configurado para enviar las características de seguridad leídas al servidor (17) a fin de determinar la identidad del participante (9), en el que la memoria de datos está dispuesta en un documento de identidad y el dispositivo de lectura (15) está configurado para leer los datos de dicha tarjeta de identidad.

15. Dispositivo para la autentificación de participantes en un servicio de telefonía que comprende lo siguiente:

un primer terminal de un participante conectado mediante un canal de transmisión a al menos un segundo terminal,

caracterizado por

una unidad de registro configurada para registrar las señales de voz transmitidas a través del canal de transmisión en el primer terminal, y
una unidad de procesamiento configurada para generar un primer conjunto de datos de verificación a partir de las señales de voz registradas,
en el que el primer terminal, además, está acoplado a un canal de seguridad a través del cual se envía un primer conjunto de datos de verificación y/o al menos un segundo conjunto de datos de verificación que se genera en el segundo terminal a partir de la señal de voz registrada en el mismo, en el que el conjunto de datos de verificación enviados a través del canal de seguridad se compara con el otro,
en el que las señales de voz registradas, a partir de las cuales se genera el primer conjunto de datos de verificación, se corresponden esencialmente con la señal de voz registrada en el segundo terminal, a partir de la cual se genera el segundo conjunto de datos de verificación.

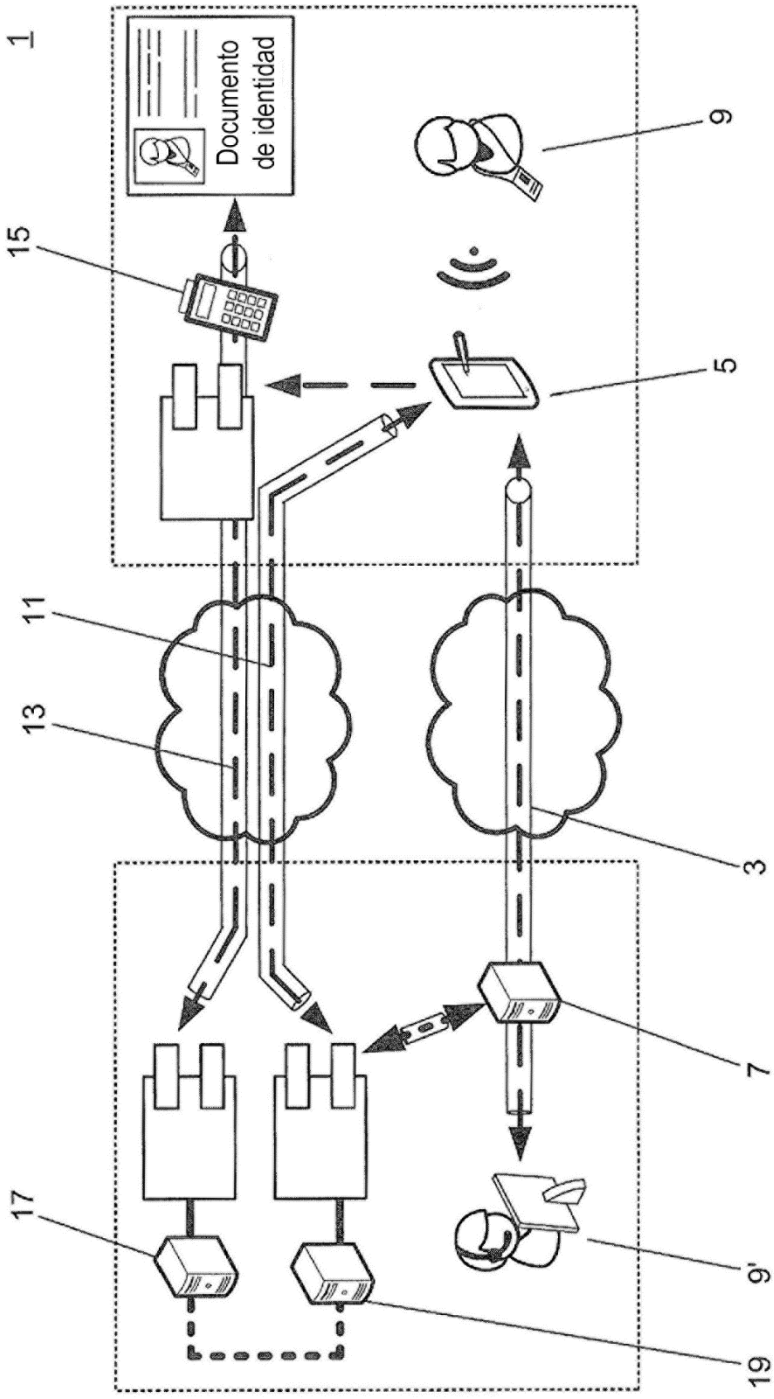

Fig. 1

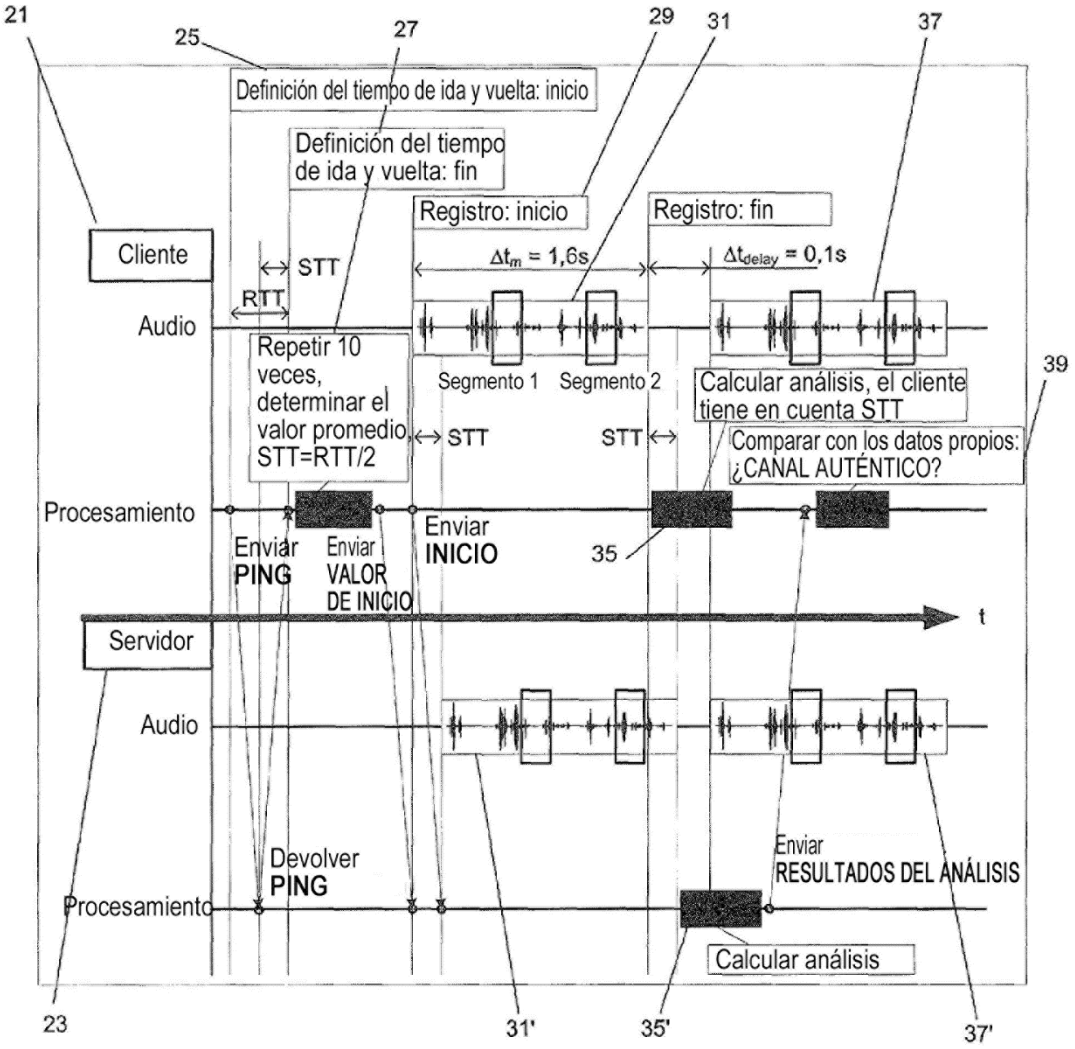
21
25
27
29
31
37
Definición del tiempo de ida y vuelta: inicio
Definición del tiempo de ida y vuelta: fin
Registro: inicio
Registro: fin
Cliente
STT
RTT
Δt_m = 1,6s
Δt_{delay} = 0,1s
Audio
Repetir 10 veces, determinar el valor promedio STT=RTT/2
Segmento 1
Segmento 2
Calcular análisis, el cliente tiene en cuenta STT
39
STT
STT
Comparar con los datos propios: ¿CANAL AUTÉNTICO?
Procesamiento
Enviar PING
Enviar VALOR DE INICIO
Enviar INICIO
35
t
Servidor
Audio
Devolver PING
Enviar RESULTADOS DEL ANÁLISIS
Procesamiento
Calcular análisis
23
31'
35'
37'

Fig. 2

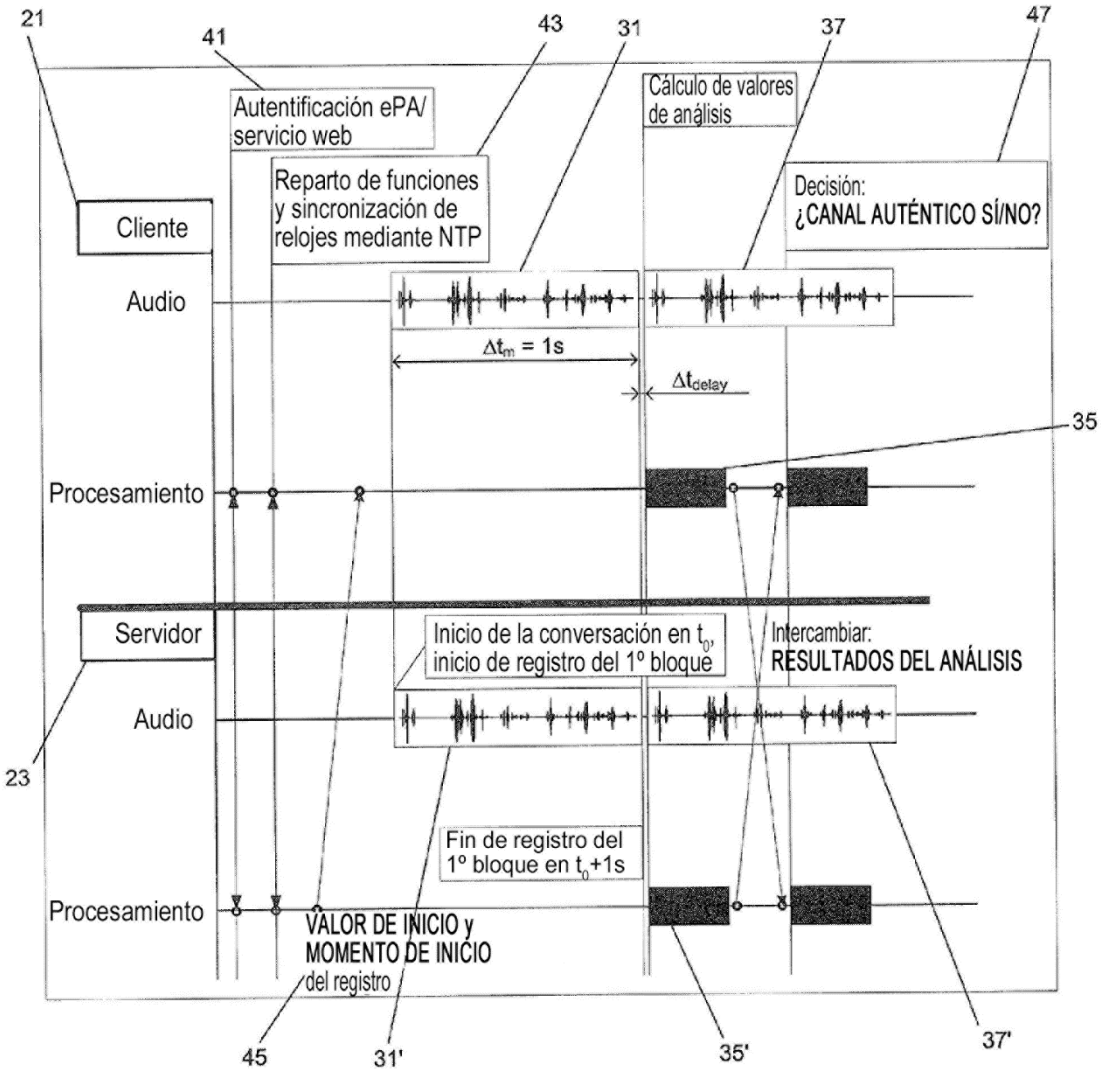
21
41
43
31
37
47
Autentificación ePA/ servicio web
Reparto de funciones y sincronización de relojes mediante NTP
Cálculo de valores de análisis
Decisión: ¿CANAL AUTÉNTICO SÍ/NO?
Cliente
Audio
Δt_m = 1s
Δt_{delay}
35
Procesamiento
Servidor
Inicio de la conversación en t_0, inicio de registro del 1º bloque
Intercambiar: RESULTADOS DEL ANÁLISIS
Audio
23
Fin de registro del 1º bloque en t_0+1s
Procesamiento
VALOR DE INICIO y MOMENTO DE INICIO del registro
45
31'
35'
37'

Fig. 3

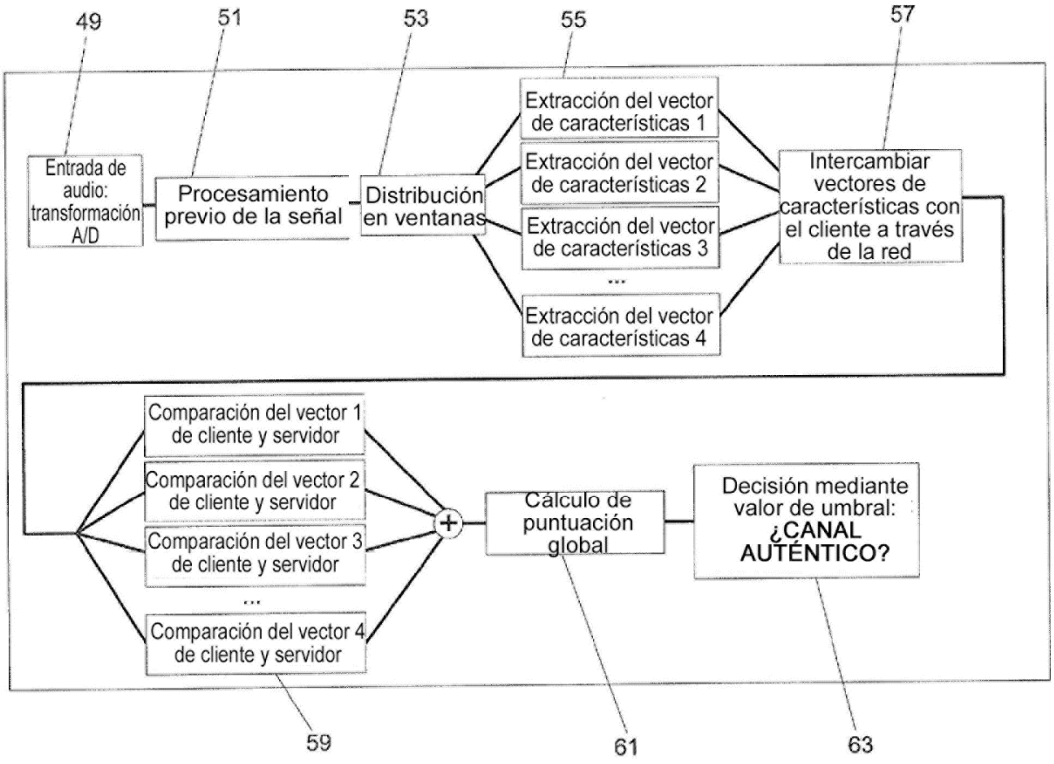
49
51
53
55
57
Entrada de audio: transformación A/D
Procesamiento previo de la señal
Distribución en ventanas
Extracción del vector de características 1
Extracción del vector de características 2
Extracción del vector de características 3
...
Extracción del vector de características 4
Intercambiar vectores de características con el cliente a través de la red
Comparación del vector 1 de cliente y servidor
Comparación del vector 2 de cliente y servidor
Comparación del vector 3 de cliente y servidor
...
Comparación del vector 4 de cliente y servidor
Cálculo de puntuación global
Decisión mediante valor de umbral: ¿CANAL AUTÉNTICO?
59
61
63

Fig. 4

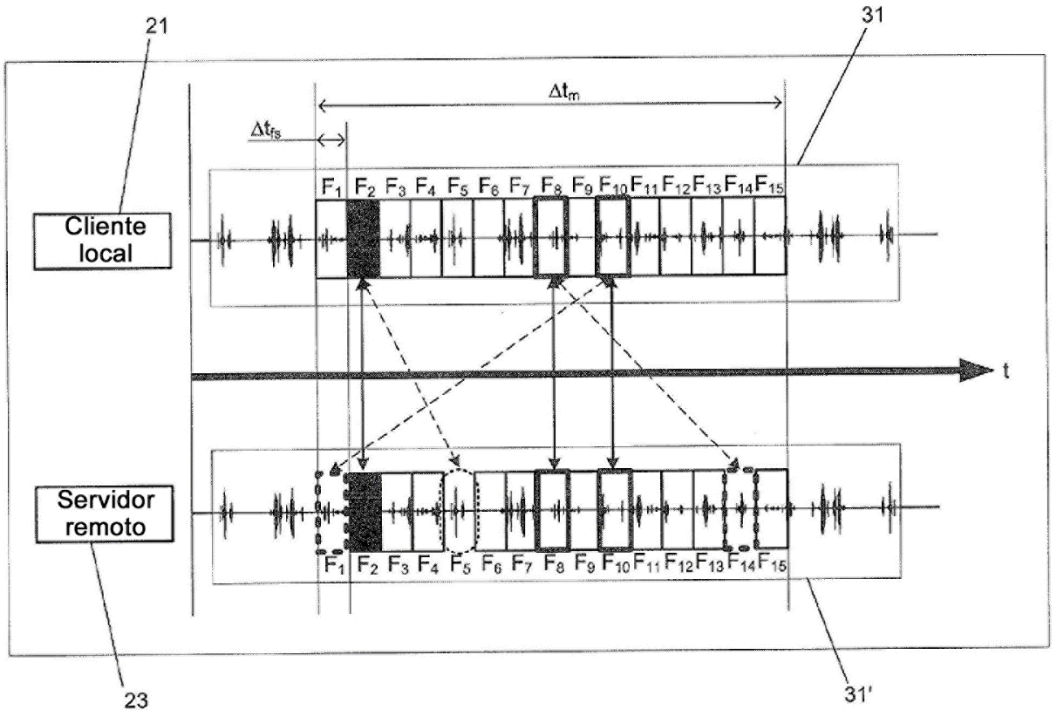
21
31
Δt_m
Δt_{fs}
F_1 F_2 F_3 F_4 F_5 F_6 F_7 F_8 F_9 F_{10} F_{11} F_{12} F_{13} F_{14} F_{15}
Cliente local
t
Servidor remoto
F_1 F_2 F_3 F_4 F_5 F_6 F_7 F_8 F_9 F_{10} F_{11} F_{12} F_{13} F_{14} F_{15}
23
31'

Fig. 5

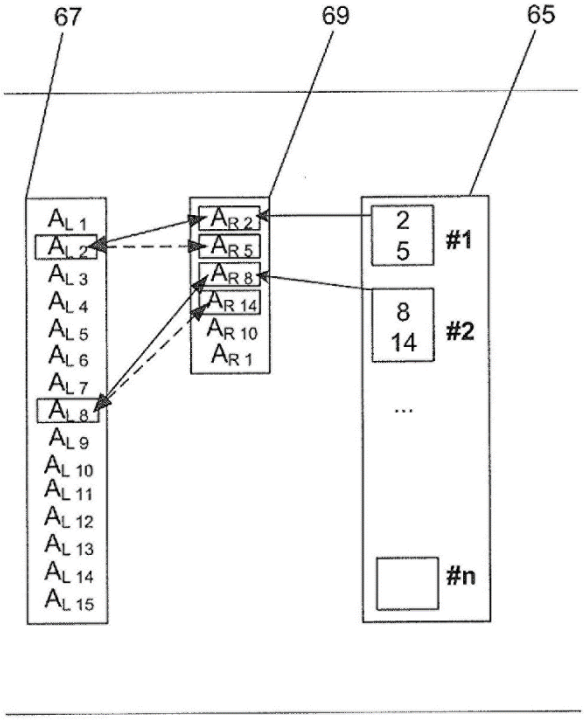
67
69
65
AL 1
AL 2
AL 3
AL 4
AL 5
AL 6
AL 7
AL 8
AL 9
AL 10
AL 11
AL 12
AL 13
AL 14
AL 15
AR 2
AR 5
AR 8
AR 14
AR 10
AR 1
2
5
#1
8
14
#2
...
#n

Fig. 6A

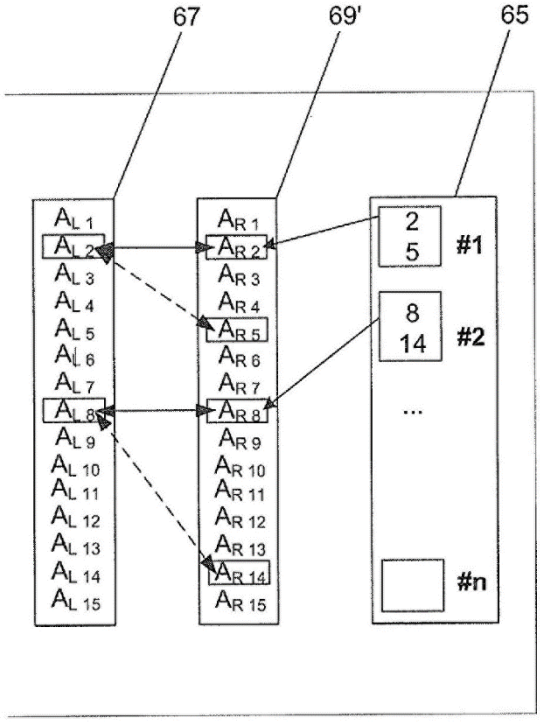
67
69'
65
AL 1
AL 2
AL 3
AL 4
AL 5
AL 6
AL 7
AL 8
AL 9
AL 10
AL 11
AL 12
AL 13
AL 14
AL 15
AR 1
AR 2
AR 3
AR 4
AR 5
AR 6
AR 7
AR 8
AR 9
AR 10
AR 11
AR 12
AR 13
AR 14
AR 15
2
5
#1
8
14
#2
...
#n

Fig. 6B

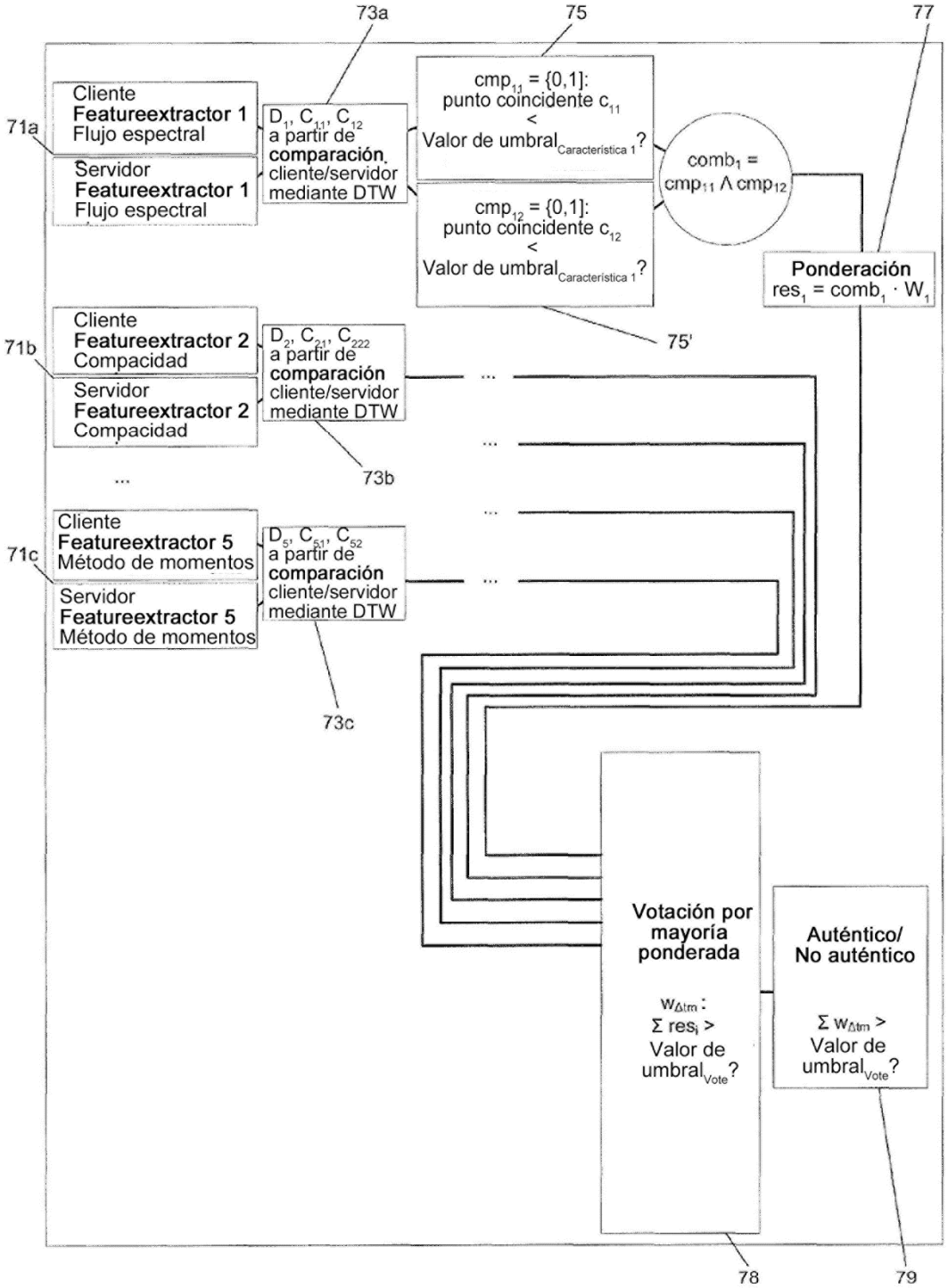
73a
75
77
Cliente
Featureextractor 1
Flujo espectral
71a
Servidor
Featureextractor 1
Flujo espectral
D1, C11, C12
a partir de
comparación
cliente/servidor
mediante DTW
cmp11 = {0,1}:
punto coincidente c11
<
Valor de umbralCaracterística 1?
cmp12 = {0,1}:
punto coincidente c12
<
Valor de umbralCaracterística 1?
comb1 =
cmp11 ∧ cmp12
Ponderación
res1 = comb1 · W1
75'
Cliente
Featureextractor 2
Compacidad
71b
Servidor
Featureextractor 2
Compacidad
D2, C211, C222
a partir de
comparación
cliente/servidor
mediante DTW
73b
...
Cliente
Featureextractor 5
Método de momentos
71c
Servidor
Featureextractor 5
Método de momentos
D5, C511, C52
a partir de
comparación
cliente/servidor
mediante DTW
73c
Votación por
mayoría
ponderada
wΔtm ·
Σ resi >
Valor de
umbralVote?
Auténtico/
No auténtico
Σ wΔtm >
Valor de
umbralVote?
78
79

Fig. 7

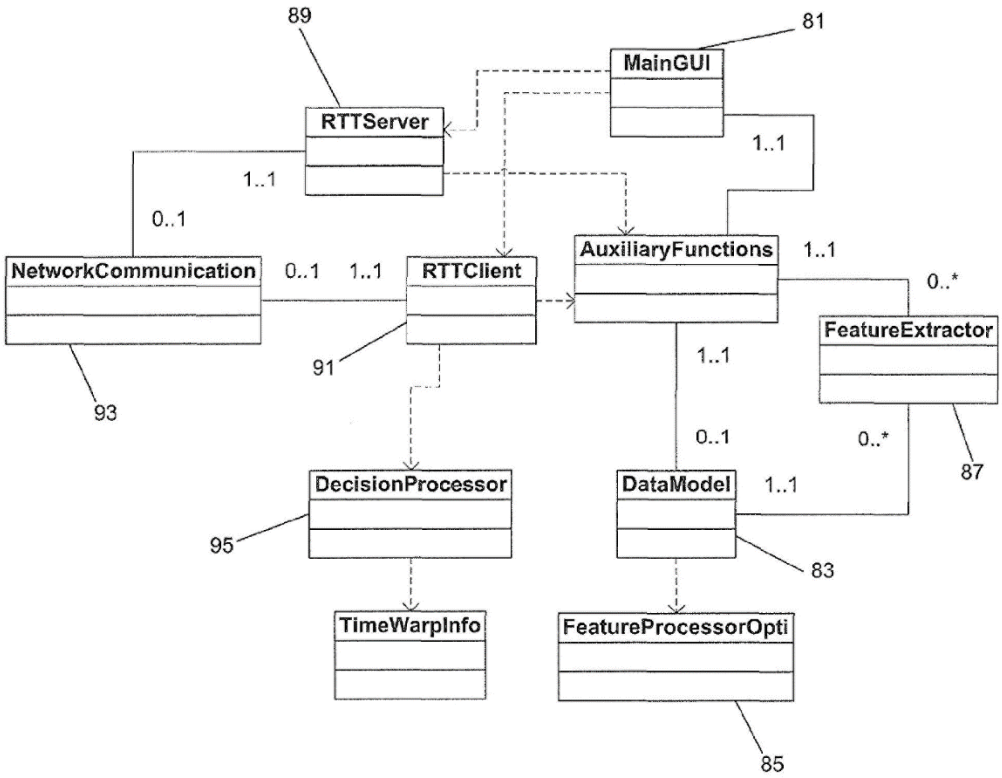
89
81
MainGUI
RTTServer
1..1
1..1
0..1
NetworkCommunication
0..1
1..1
RTTClient
AuxiliaryFunctions
1..1
0..*
FeatureExtractor
91
1..1
93
0..1
0..*
87
DecisionProcessor
DataModel
1..1
95
83
TimeWarpInfo
FeatureProcessorOpti
85

Fig. 8