

UIT-T

SECTOR DE NORMALIZACIÓN
DE LAS TELECOMUNICACIONES
DE LA UIT

E.840

(06/2018)

SERIE E: EXPLOTACIÓN GENERAL DE LA RED, SERVICIO TELEFÓNICO, EXPLOTACIÓN DEL SERVICIO Y FACTORES HUMANOS

Calidad de los servicios de telecomunicación: conceptos, modelos, objetivos, planificación de la seguridad de funcionamiento – Modelos para los servicios de telecomunicación

Marco estadístico de referencia para la puntuación y calificación de la calidad de funcionamiento de la red de extremo a extremo

Recomendación UIT-T E.840

RECOMENDACIONES UIT-T DE LA SERIE E

EXPLOTACIÓN GENERAL DE LA RED, SERVICIO TELEFÓNICO, EXPLOTACIÓN DEL SERVICIO Y FACTORES HUMANOS

EXPLOTACIÓN DE LAS RELACIONES INTERNACIONALES	
Definiciones	E.100–E.103
Disposiciones de carácter general relativas a las Administraciones	E.104–E.119
Disposiciones de carácter general relativas a los usuarios	E.120–E.139
Explotación de las relaciones telefónicas internacionales	E.140–E.159
Plan de numeración del servicio telefónico internacional	E.160–E.169
Plan de encaminamiento internacional	E.170–E.179
Tonos utilizados en los sistemas nacionales de señalización	E.180–E.189
Plan de numeración del servicio telefónico internacional	E.190–E.199
Servicio móvil marítimo y servicio móvil terrestre público	E.200–E.229
DISPOSICIONES OPERACIONALES RELATIVAS A LA TASACIÓN Y A LA CONTABILIDAD EN EL SERVICIO TELEFÓNICO INTERNACIONAL	
Tasación en el servicio internacional	E.230–E.249
Medidas y registro de la duración de las conferencias a efectos de la contabilidad	E.260–E.269
UTILIZACIÓN DE LA RED TELEFÓNICA INTERNACIONAL PARA APLICACIONES NO TELEFÓNICAS	
Generalidades	E.300–E.319
Telefotografía	E.320–E.329
DISPOSICIONES DE LA RDSI RELATIVAS A LOS USUARIOS	
PLAN DE ENCAMINAMIENTO INTERNACIONAL	
GESTIÓN DE RED	
Estadísticas relativas al servicio internacional	E.400–E.404
Gestión de la red internacional	E.405–E.419
Comprobación de la calidad del servicio telefónico internacional	E.420–E.489
INGENIERÍA DE TRÁFICO	
Medidas y registro del tráfico	E.490–E.505
Previsiones del tráfico	E.506–E.509
Determinación del número de circuitos necesarios en explotación manual	E.510–E.519
Determinación del número de circuitos necesarios en explotación automática y semiautomática	E.520–E.539
Grado de servicio	E.540–E.599
Definiciones	E.600–E.649
Ingeniería de tráfico para redes con protocolo Internet	E.650–E.699
Ingeniería de tráfico de RDSI	E.700–E.749
Ingeniería de tráfico de redes móviles	E.750–E.799
CALIDAD DE LOS SERVICIOS DE TELECOMUNICACIÓN: CONCEPTOS, MODELOS, OBJETIVOS, PLANIFICACIÓN DE LA SEGURIDAD DE FUNCIONAMIENTO	
Términos y definiciones relativos a la calidad de los servicios de telecomunicación	E.800–E.809
Modelos para los servicios de telecomunicación	E.810–E.844
Objetivos para la calidad de servicio y conceptos conexos de los servicios de telecomunicaciones	E.845–E.859
Utilización de los objetivos de calidad de servicio para la planificación de redes de telecomunicaciones.	E.860–E.879
Recopilación y evaluación de datos reales sobre la calidad de funcionamiento de equipos, redes y servicios	E.880–E.899
OTROS	E.900–E.999
EXPLOTACIÓN DE LAS RELACIONES INTERNACIONALES	
Plan de numeración del servicio telefónico internacional	E.1100–E.1199
GESTIÓN DE LAS REDES	
Gestión de las redes internacionales	E.4100–E.4199

Para más información, véase la Lista de Recomendaciones del UIT-T.

Recomendación UIT-T E.840

Marco estadístico de referencia para la puntuación y la calificación de la calidad de funcionamiento de la red de extremo a extremo

Resumen

La Recomendación UIT-T E.840 es la primera de una serie de recomendaciones sobre la determinación de los criterios básicos de la calidad de funcionamiento de la red de extremo a extremo. En la Recomendación UIT-T E.840 se presenta un marco para el análisis estadístico subyacente a la fijación de criterios de calidad de funcionamiento de redes y servicios. En el marco se describen las hipótesis, los casos de uso, los procedimientos y técnicas estadísticas de referencia para establecer una clasificación de los indicadores fundamentales de rendimiento (IFR) o indicadores fundamentales de calidad (IFC) de la calidad de funcionamiento de extremo a extremo. La Recomendación UIT-T E.840 atañe a los servicios móviles y las campañas de determinación de criterios realizadas mediante agentes (dispositivos) móviles en pruebas realizadas durante paseos en vehículos o a pie y agentes fijos, es decir, dispositivos situados en emplazamientos fijos (por ejemplo, centros comerciales, oficinas o estadios).

Historia

Edición	Recomendación	Aprobación	Comisión de Estudio	ID único*
1.0	ITU-T E.840	2018-06-13	12	11.1002/1000/13621

Palabras clave

Calidad de funcionamiento de extremo a extremo, determinación de criterios básicos y clasificación de la calidad de funcionamiento de red, marco estadístico.

* Para acceder a la Recomendación, sírvase digitar el URL <http://handle.itu.int/> en el campo de dirección del navegador, seguido por el identificador único de la Recomendación. Por ejemplo, <http://handle.itu.int/11.1002/1000/11830-en>.

PREFACIO

La Unión Internacional de Telecomunicaciones (UIT) es el organismo especializado de las Naciones Unidas en el campo de las telecomunicaciones y de las tecnologías de la información y la comunicación. El Sector de Normalización de las Telecomunicaciones de la UIT (UIT-T) es un órgano permanente de la UIT. Este órgano estudia los aspectos técnicos, de explotación y tarifarios y publica Recomendaciones sobre los mismos, con miras a la normalización de las telecomunicaciones en el plano mundial.

La Asamblea Mundial de Normalización de las Telecomunicaciones (AMNT), que se celebra cada cuatro años, establece los temas que han de estudiar las Comisiones de Estudio del UIT-T, que a su vez producen Recomendaciones sobre dichos temas.

La aprobación de Recomendaciones por los Miembros del UIT-T es el objeto del procedimiento establecido en la Resolución 1 de la AMNT.

En ciertos sectores de la tecnología de la información que corresponden a la esfera de competencia del UIT-T, se preparan las normas necesarias en colaboración con la ISO y la CEI.

NOTA

En esta Recomendación, la expresión "Administración" se utiliza para designar, en forma abreviada, tanto una administración de telecomunicaciones como una empresa de explotación reconocida de telecomunicaciones.

La observancia de esta Recomendación es voluntaria. Ahora bien, la Recomendación puede contener ciertas disposiciones obligatorias (para asegurar, por ejemplo, la aplicabilidad o la interoperabilidad), por lo que la observancia se consigue con el cumplimiento exacto y puntual de todas las disposiciones obligatorias. La obligatoriedad de un elemento preceptivo o requisito se expresa mediante las frases "tener que, haber de, hay que + infinitivo" o el verbo principal en tiempo futuro simple de mandato, en modo afirmativo o negativo. El hecho de que se utilice esta formulación no entraña que la observancia se imponga a ninguna de las partes.

PROPIEDAD INTELECTUAL

La UIT señala a la atención la posibilidad de que la utilización o aplicación de la presente Recomendación suponga el empleo de un derecho de propiedad intelectual reivindicado. La UIT no adopta ninguna posición en cuanto a la demostración, validez o aplicabilidad de los derechos de propiedad intelectual reivindicados, ya sea por los miembros de la UIT o por terceros ajenos al proceso de elaboración de Recomendaciones.

En la fecha de aprobación de la presente Recomendación, la UIT no ha recibido notificación de propiedad intelectual, protegida por patente, que puede ser necesaria para aplicar esta Recomendación. Sin embargo, debe señalarse a los usuarios que puede que esta información no se encuentre totalmente actualizada al respecto, por lo que se les insta encarecidamente a consultar la base de datos sobre patentes de la TSB en la dirección <http://www.itu.int/ITU-T/ipr/>.

© UIT 2020

Reservados todos los derechos. Ninguna parte de esta publicación puede reproducirse por ningún procedimiento sin previa autorización escrita por parte de la UIT.

ÍNDICE

	Página
1 Alcance	1
2 Referencias	1
3 Definiciones.....	1
4 Abreviaturas y acrónimos	1
5 Convenios	2
6 Tipos de comparación.....	2
7 Condiciones de comparación.....	3
8 Servicios comparados	4
9 Marco estadístico	5
9.1 Depuración de datos	5
9.2 Distribución estadística de la medición.....	5
9.3 Unidad de calidad de funcionamiento estadística, error típico y significación estadística de los resultados de la comparación.....	6
9.4 Puntuación y clasificación de IFR o IFC de extremo a extremo.....	7
Anexo A – Aplicación de la significación estadística al análisis comparativo de redes móviles.....	9
Anexo B – Puntuación y clasificación estadísticas de la calidad de funcionamiento de una red	10
Apéndice I – Posible técnica para la puntuación y clasificación estadísticas de una red	12
Bibliografía	14

Recomendación UIT-T E.840

Marco estadístico de referencia para la puntuación y la calificación de la calidad de funcionamiento de la red de extremo a extremo

1 Alcance

En esta Recomendación se especifica un marco estadístico, además de las hipótesis y condiciones de referencia en que puede aplicarse, que los operadores y reguladores han de utilizar a la hora de calificar y cuantificar las diferencias en calidad de funcionamiento de extremo a extremo en función de los indicadores fundamentales de rendimiento (IFR) y los indicadores fundamentales de calidad (IFC) relativos a la experiencia del usuario.

La necesidad de esta Recomendación surge porque, en su empeño por satisfacer las expectativas de usuarios cada vez más exigentes al tiempo que amplían su base de clientes a un coste óptimo, los operadores han mejorado la calidad de funcionamiento de las redes hasta tal punto que las diferencias entre ellos son cada vez más insignificantes.

2 Referencias

Las siguientes Recomendaciones UIT-T y demás referencias contienen disposiciones que, por referencia a las mismas en este texto, constituyen disposiciones de esta Recomendación. En la fecha de publicación, las ediciones citadas estaban en vigor. Todas las Recomendaciones y demás referencias están sujetas a revisión, por lo que se alienta a los usuarios de esta Recomendación a que consideren la posibilidad de aplicar la edición más reciente de las Recomendaciones y demás referencias que se indican a continuación. Se publica periódicamente una lista de las Recomendaciones UIT-T vigentes. La referencia a un documento en el marco de esta Recomendación no confiere al mismo, como documento autónomo, el rango de Recomendación.

[UIT-T E.800] Recomendación UIT-T E.800 (2008), *Definiciones de los términos relativos a la calidad de servicio*.

[UIT-T E.804] Recomendación UIT-T E.804 (2014), *Aspectos relativos a la calidad de servicio para los servicios más utilizados en las redes móviles*.

3 Definiciones

Ninguna.

4 Abreviaturas y acrónimos

En esta Recomendación se utilizan las siguientes abreviaturas y acrónimos:

IFR Indicador fundamental de rendimiento

IFC Indicador fundamental de calidad

QoE Calidad percibida (*quality of experience*)

MOS Nota media de opinión (*mean opinion score*)

RF Radiofrecuencia

TCP Protocolo de control de transmisión (*transmission control protocol*)

5 Convenios

5.2.1 StatScore: Representa una nota estadística, es decir, la calidad general relativa entre distintas redes u operadores en comparación con la red con mejor calidad de funcionamiento. StatScore se calcula para cada servicio.

5.2.2 GlobalNetScore: Representa una nota de red global, es decir, la calidad general relativa entre distintas redes u operadores en comparación con la red con mejor calidad de funcionamiento. GlobalNetScore se calcula para todos los servicios.

5.2.3 StatDiff: Representa la diferencia estadísticamente importante al comparar dos indicadores fundamentales de rendimiento (IFR) o indicadores fundamentales de calidad (IFC).

5.2.4 THrelv: Representa la diferencia mínima entre los valores de dos indicadores fundamentales de rendimiento (IFR) o indicadores fundamentales de calidad (IFC), pertinente para el usuario de un servicio y sobre la que la significación estadística es prioritaria.

6 Tipos de comparación

Por norma general, hay dos subtipos de comparación de redes: interna y competitiva. La comparación interna se centra en garantizar una calidad de funcionamiento de red constante y rentable y aportar mejoras para las que es necesario hacer evaluaciones al inicio del despliegue de la red, así como durante su desarrollo y cuando se lanzan nuevos servicios y dispositivos. La comparación interna también se realiza en redes maduras e implantadas. Además, durante las campañas de evaluación se han de tener en cuenta las regiones con autopistas y ciudades, así como las zonas de interés (por ejemplo, lugares de trabajo, centros comerciales, estadios e inmuebles residenciales). La comparación competitiva la realizan los operadores mismos (o empresas terceras en su nombre), así como los reguladores a la hora de verificar la competencia. La autoclasificación suele llevarse a cabo en regiones, rutas de transporte (autopistas, ferrocarriles) y ciudades, e incluso países, en el caso de operadores multinacionales, y en redes maduras.

En la Figura 1 se resumen los casos de uso y las herramientas y técnicas recomendadas. Las zonas (como los centros comerciales, los estados y lugares de trabajo, generalmente en interiores) se suelen comparar a la prueba de paseo. Además de las habituales campañas de prueba mediante paseos, a pie o en automóviles, para las comparaciones internas y, en cierta medida también para las competitivas, y las evaluaciones en interiores también se utilizan herramientas fijas tipo sonda. Estas tienen la ventaja de poder adaptarse rápidamente y a distancia, además de su independencia de los dispositivos. Por consiguiente, estas herramientas se adaptan muy bien a las pruebas en interiores y al lanzamiento de nuevos servicios en zonas de interés, aunque también en ciudades. Además, como puede verse, se pueden aplicar técnicas de análisis *a posteriori* o *a priori*. Con las primeras, generalmente utilizadas para determinar criterios básicos, se obtienen los datos y se emplea la significación estadística para evaluar y clasificar la calidad de funcionamiento de acuerdo con IFR o IFC de extremo a extremo. La precisión de la medición está integrada por defecto en el nivel de significación estadística. Las técnicas *a priori* exigen el cálculo previo del número de sondas de prueba necesarias para una significación estadística y una precisión de medición específicas. Estas técnicas suelen emplearse cuando las sondas son onerosas o el tiempo de prueba es limitado.

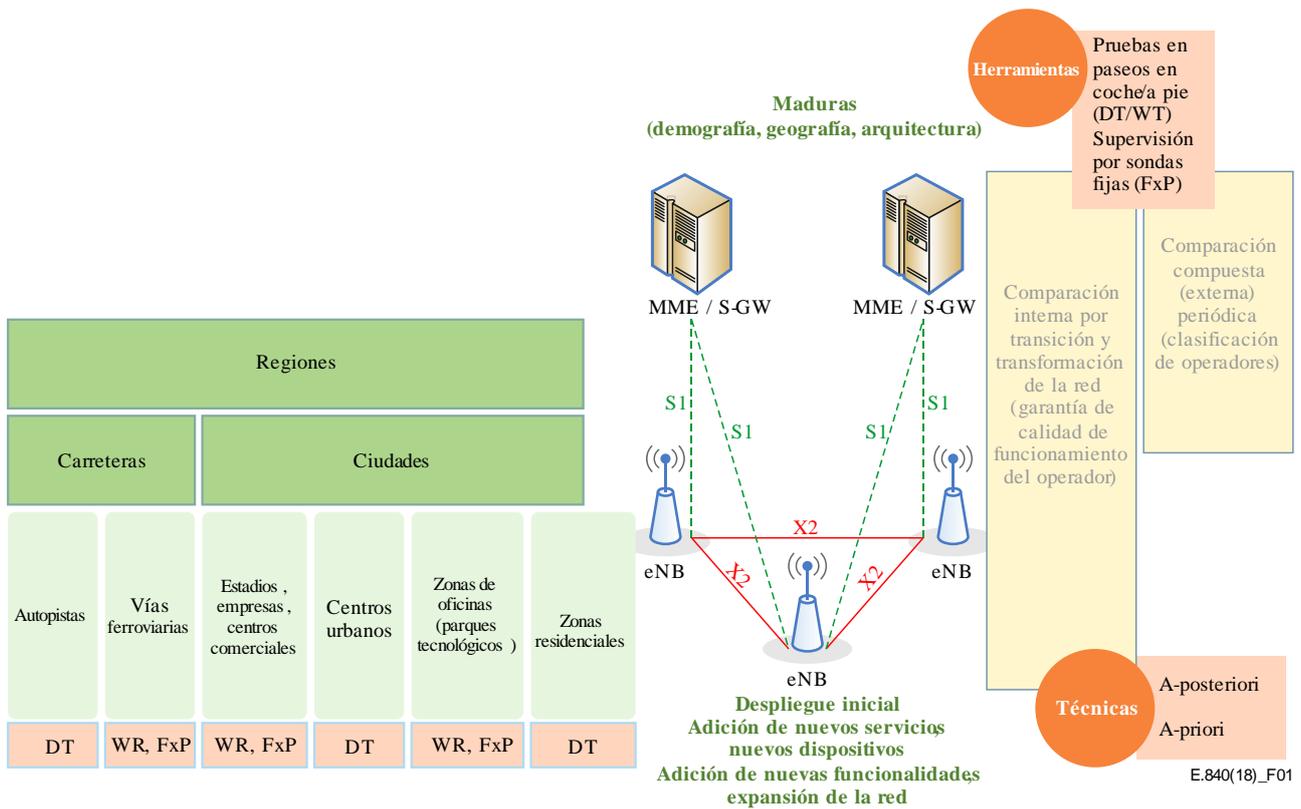


Figura 1 – Casos de uso, herramientas y técnicas para la comparación

7 Condiciones de comparación

Sea cual sea el uso a que está destinado, el marco de comparación debe basarse en una serie de requisitos previos que garanticen la coherencia, la validez, la fiabilidad y la repetibilidad. En el Cuadro 1 se muestran los requisitos previos de cada fase de la comparación: configuración del equipo, configuración del aprueba, obtención de datos, procesamiento y análisis de datos. Cabe señalar que en el Cuadro 1 se indican los requisitos previos mínimos necesarios para garantizar el control total del entorno de prueba, así como la validez del análisis estadístico. Los detalles concretos de las mediciones pueden encontrarse en otras Recomendaciones UIT-T (por ejemplo, [UIT-T E.804]).

Cuadro 1 – Resumen de los requisitos previos mínimos necesarios

Fase de la comparación	Requisitos previos
Configuración del equipo	La configuración del equipo ha de ser coherente entre redes, plataformas y dispositivos. Se han de utilizar el mismo modelo de dispositivo para la determinación de criterios básicos <i>competitiva comparativa</i> ("entre iguales").
	El equipo debe funcionar en las condiciones adecuadas descritas por el fabricante del equipo de prueba (por ejemplo, hay que evitar el sobrecalentamiento, pues puede menoscabar la calidad de funcionamiento del dispositivo).
Configuración de la prueba	Se debe configurar la prueba para reflejar la experiencia real del usuario. Pueden encontrarse orientaciones al respecto en [b-ETSI TR 102 581].
	Los parámetros de configuración de la prueba deben evitar artefactos que puedan modificar artificialmente la calidad de funcionamiento de la red (por ejemplo, se han de verificar la ubicación y configuración del servidor de datos mediante los parámetros del protocolo de control de transmisión (TCP) para garantizar un buen caudal para todos los operadores objeto de la comparación). Esta prueba se ha de realizar antes de proceder a la obtención de los datos.
	La prueba se debe configurar teniendo en cuenta las diversas hipótesis de comportamiento del usuario (por ejemplo, utilización de TCP, ficheros o llamadas de vídeo o voz de diversa longitud), minimizando al mismo tiempo la intervención para no sobrecargar artificialmente la red.
Obtención de datos	Se deben recopilar los datos de las mediciones que reflejen la experiencia del usuario (por ejemplo, nota media de opinión (MOS) por servicio y los IFR o IFC de extremo a extremo de la red principal que influyan en él). Para las mediciones deben utilizarse las unidades o técnicas de medición adecuadas de conformidad con las directrices del fabricante o de las Recomendaciones pertinentes.
	Deben obtenerse los datos de las mediciones efectuadas con diversas condiciones geográficas o demográficas y en distintos momentos (hora punta o baja, fin de semana o día laborable, periodo de vacaciones o de trabajo).
Procesamiento y análisis de datos	Deben compararse los datos obtenidos con los mismos dispositivos en la misma zona y durante el mismo periodo de tiempo – comparación "entre iguales".
	Debe utilizarse la significación estadística para efectuar comparaciones significativas.
	Se ha de realizar un análisis en función de los IFR o los IFC

8 Servicios comparados

Quedan fuera del alcance de esta Recomendación la lista de servicios móviles habitualmente comparados, así como sus IFR o IFC y los correspondientes valores de referencia. Pueden obtenerse más información al respecto en [UIT-T E.800] y [UIT-T E.804].

Si el objetivo de la comparación móvil es realizar un análisis comparativo detallado por servicios, cosa que suele realizarse mediante la comparación interna (por ejemplo, introducción de nuevos dispositivos, nuevas tecnologías), se recomienda utilizar para el análisis un conjunto completo de IFR e IFC (aquí los IFC son mediciones obtenidas utilizando modelos de estimación de la calidad, como [b-UIT-T P.863] para la voz o [b-UIT-T P.1203] para la difusión de vídeo). Además, se recomienda analizar con los mismos criterios las principales causas que pueden motivar una baja calidad de funcionamiento.

Por otra parte, si el objetivo de la comparación móvil es la clasificación en función de IFR o IFC de calidad de funcionamiento de extremo a extremo, lo que se suele realizar mediante comparación competitiva, aunque también interna en algunos casos (como las comparaciones de mercado y la evaluación periódica del rendimiento de mercado), para la puntuación y clasificación se puede considerar un conjunto más reducido de IFR o IFC que influyan en la calidad percibida (QoE, véase [b-UIT-T P.10/G.100]) por servicios y en todos los servicios objeto de la comparación.

En esta Recomendación se contempla este último caso de comparación competitiva. En otras Recomendaciones UIT-T sobre la comparación de referencia se detallan los conjuntos de IFR o IFC.

9 Marco estadístico

El objetivo del marco recomendado es puntuar y clasificar la calidad de funcionamiento de extremo a extremo de la red desde la perspectiva del usuario. Este marco puede utilizarse para comparaciones competitivas o internas. En el marco se definen los procedimientos para la validación de los datos, las unidades de evaluación estadística y las pruebas de significación, así como las directrices generales para la puntuación y la clasificación.

9.1 Depuración de datos

Para garantizar que los resultados de la comparación son significativos y precisos es necesario validar los datos utilizados para el análisis. Esta validación consiste básicamente en una depuración de los datos, es decir, la eliminación de todos los datos incompletos o factores propios de la medición. Se recomienda sustituir los datos eliminados con los datos de una nueva medición. Si el 5% o más de los datos contiene artefactos o se ha eliminado, se recomienda volver a tomar todos los datos. Los artefactos propios de la medición pueden ser, entre otros, las llamadas silentes en el caso de los servicios de conversación de voz o vídeo, los periodos de indisponibilidad del servidor de datos o vídeo (por ejemplo, caída del servidor) en el caso de los servicios de difusión de vídeo, o valores MOS constante e inesperadamente bajos para la voz, la videoconversación o la difusión. Toda degradación no causada por la red misma, sino por los equipos o dispositivos de prueba, se ha de considerar como un artefacto propio de la medición.

Es necesario filtrar los datos de la comparación y eliminar esos artefactos para garantizar la validez de los supuestos al aplicar los modelos estadísticos.

9.2 Distribución estadística de la medición

Por norma general, la distribución estadística de los valores de cualquier IFR o IFC medido puede aproximarse mediante una distribución gaussiana sobre la base del teorema del límite central [b-Shaum]: cuanto mayor sea el número de muestras, mayor será la precisión de la aproximación a la distribución gaussiana. Se recomienda verificar la distribución estadística de todos los valores de IFR o IFC, lo que puede hacerse de dos maneras: pueden generarse diagramas de distribución de los valores medidos para los IFR o IFC considerados y verificar su normalidad mediante una simple observación, o puede verificarse la normalidad mediante pruebas de adecuación como las de Kolmogorov-Smirnov, Anderson-Darling o Shapiro-Wilk [b-Mehta]. Además, en el caso raro y extremo de que la distribución no sea gaussiana, pueden efectuarse pruebas no paramétricas. Las pruebas que dan distribuciones experimentales de diversos IFR o IFC, las pruebas de verificación de la normalidad de la distribución y los casos especiales de distribuciones no gaussianas se abordan en otras Recomendaciones UIT-T de la serie de comparaciones de referencia.

9.3 Unidad de calidad de funcionamiento estadística, error típico y significación estadística de los resultados de la comparación

9.3.1 Unidad de calidad de funcionamiento estadística

Los análisis comparativos deben basarse en unidades de calidad de funcionamiento estadísticas que reflejen la calidad de funcionamiento media de la red (representada por valores medios, m) o su coherencia (representada por la P-ésima probabilidad de estar por encima de un valor umbral predefinido). En esta Recomendación se utiliza como ejemplo la calidad de funcionamiento estadística media. Otras técnicas semejantes pueden aplicarse para la coherencia.

9.3.2 Error típico

El error típico con un nivel de confianza del 95% para la media o la P-ésima se calcula suponiendo que los IFR o IFC medidos tienen una distribución gaussiana (véase la cláusula 9.2).

Por consiguiente, en función del tipo de IFR o IFC de que se trate, puntuaciones continuas como las del parámetro radiofrecuencia (RF) o MOS o puntuaciones discretas como las de la relación éxito-fallo (r), el error típico con un nivel de confianza del 95% se obtiene así:

$$\text{StdError}(m) = z_{95\%} * \text{std} / \sqrt{N} = 1,96 * \text{std} / \sqrt{N}$$

$$\text{StdError}(r) = z_{95\%} * \sqrt{r * (1-r) / N} = 1,96 * \sqrt{r * (1-r) / N}$$

Si se dispone de menos de 30 muestras, deberá sustituirse el cuantil gaussiano $z_{95\%}$ por el valor tabulado Student $t_{95\%}$ ($N-1$), siendo N el número de muestras disponibles.

NOTA – El error típico representa la precisión de la medición. Por consiguiente, si se necesita una precisión concreta y se conoce la estimación de la desviación típica, puede determinarse el número mínimo de muestras necesarias para alcanzar esa precisión con un nivel de confianza seleccionado utilizando las ecuaciones del párrafo 2. Esto puede utilizarse con las técnicas *a priori* de la Figura 1 y también en [b-UIT-T E.802] para calcular el número mínimo de muestras.

9.3.3 Significación estadística

Los intervalos de confianza son los márgenes de error típico para un determinado nivel de confianza estadística, generalmente el 95%. Sin embargo, dos IFR o IFC comparados con valores cercanos e intervalos de confianza coincidentes no necesariamente son estadísticamente iguales. Una comparación precisa entre IFR o IFC debe basarse en la significación estadística, que garantiza que la probabilidad de rechazar falsamente la hipótesis de que dos valores de IFR o IFC son iguales, cuando efectivamente son iguales, no supera el 5%.

Gracias a la prueba de significación estadística puede llegarse a la conclusión de que un operador ofrece una calidad de funcionamiento mejor que otro (en el caso de la comparación competitiva) y determinar si una nueva tecnología o funcionalidad aporta una mejora consecuente (en el caso de la comparación interna).

Además de la significación estadística, se han de utilizar los umbrales de diferencia específicos de IFR o IFC pertinentes cuando las diferencias son insignificantes o dentro del margen de precisión de la medición de cada IFR o IFC. En otras Recomendaciones UIT-T de la serie sobre comparación de referencia se definen los umbrales de diferencia específicos de IFR o IFC pertinentes (TH_{relv}).

En el Cuadro 2 se presenta un ejemplo de cómo se han de comparar los IFR o IFC. Para mantener el carácter general de esta Recomendación, los elementos que se comparan son IFC1 e IFC2. Por consiguiente, IFC1 e IFC2 pueden ser cualquier IFC escogido para un servicio concreto (también puede ser un IFR). Se calculan los valores de desviación media y típica para IFC1 e IFC2. Aplicando las Ecuaciones A-1 y A-2 con el número de pruebas disponibles, se obtiene que la calidad ofrecida por las redes 1 y 2 en lo que respecta a IFC1 e IFC2 es estadísticamente idéntica. Además, puede verse que las diferencias para cada IFC, IFC1 e IFC2, es inferior a TH_{relv} . La calidad de funcionamiento de las redes 1 y 2 es estadísticamente diferente cuando se considera una

tercera medida, IFC3. Sin embargo, la diferencia (0,02) permanece por debajo de THrelv (0,025), por lo que no puede concluirse que la calidad de funcionamiento de las redes 1 y 2 sea diferente.

Cuadro 2 – Ejemplo de significación estadística

Servicio	IFC	Red 1			Red 2			Nivel de confianza estadística del 95% (Z>1,96)	Diferencia estadística	THrelv
		Media	Típica	N	Media	Típica	N			
Service 1	IFC1	3,27	0,3	287	3,35	0,6	212	1,78	No	0,09
	IFC2	0,02	0,14	12	0,015	0,12	10	0,09	No	0,006
	IFC3	0,93	0,26	69	0,91	0,29	71	2,04	Sí	0,025

Este tipo de análisis aplicado a los resultados de una comparación detallada puede ampliarse a varios servicios, así como a un conjunto más amplio de IFR o IFC para un servicio concreto, como se indica en la cláusula 8. Además, en función de los resultados de significación estadística (Z, nivel de confianza estadística del 95%, Cuadro 2), puede establecerse una clasificación de IFR o IFC entre redes, como se indica en la cláusula 9.4.

Cabe señalar que para afirmar que una red o configuración puede considerarse "mejor" que otra se necesita, además de la significación estadística, una definición del umbral de diferencia específicos de IFR o IFC pertinentes e información sobre la precisión de la medición, que se definen en otras Recomendaciones UIT-T de la serie sobre comparación de referencia.

9.3.4 Rendición de resultados

El análisis estadístico comparativo de referencia y sus resultados han de presentarse junto con una descripción detallada de la configuración y condiciones de las pruebas utilizadas para la comparación. En caso contrario los resultados podrán interpretarse erróneamente y perder su relevancia.

9.4 Puntuación y clasificación de IFR o IFC de extremo a extremo

Para cada una de las esferas consideradas en la campaña de comparación puede establecerse una puntuación o clasificación de los IFR o IFC de la red de extremo a extremo. Además, como ya se indicó en la cláusula 9.3.3, es necesario basarse en la significación estadística para detectar diferencias mínimas en la calidad de funcionamiento, lo que ocurre con cada vez mayor frecuencia en las actuales redes de operador. Se ha de recurrir a la comparación de los umbrales de diferencia específicos de IFR o IFC cuando las diferencias son insignificantes o se sitúan dentro del margen de precisión de las mediciones, como se indica en otras Recomendaciones UIT-T de la serie dedicada a la comparación de referencia.

En el Anexo B se presenta la metodología de puntuación y clasificación estadística para un conjunto de IFR o IFC. Un ejemplo de ello se muestra en el Cuadro 3 siguiente.

Cuadro 3 – Ejemplo de puntuación y clasificación de IFR o IFC de extremo a extremo con significación estadística

	Red 1				Red 2				THrelv
	IFR/IFC	Típica	N	StatDiff	IFR/IFC	Típica	N	StatDiff	
IFR1/IFC1	0,95	0,22	87	0,05	0,97	0,17	69	0,00	0,018
IFR2/IFC2	0,93	0,26	87	0,00	0,91	0,29	69	0,23	0,019
IFR3/IFC3	3,89	0,50	2 600	0,00	3,56	0,70	2 070	17,15	0,31
IFR4/IFC4	105,00	5,00	435	42,67	70,00	15,00	350	0,00	34
IFR5/IFC5	1 200,00	300,00	87	0,00	1 800,00	275,00	69	12,31	596

En el Cuadro 3 se muestran los valores combinados correspondientes a la misma esfera durante el mismo periodo de tiempo para una comparación "entre iguales" (véase el Cuadro 1). Además de los valores de calidad de funcionamiento, se calculan las desviaciones típicas y se indica el número de muestras. Se calcula StatDiff (como se indica en el Anexo B) para cada IFR o IFC como diferencia estadísticamente significativa con respecto al mejor IFR o IFC (resaltado en amarillo en el Cuadro 3) con un nivel de confianza del 95% (véase el Anexo B). Cuanto menor sea el valor de StatDiff, más cercano estará el valor al mejor IFR o IFC. Un StatDiff = 0 representa el mejor IFR o IFC. Además, cabe señalar que, en todos los casos, las diferencias entre IFC correspondientes a las dos redes son superiores a THrevl, lo que implica que el análisis de significación estadística es prioritario.

De acuerdo con este análisis, en el Cuadro 4 se muestra la clasificación de IFR o IFC.

Cuadro 4 – Ejemplo de clasificación estadística de IFR o IFC

IFR/IFC	Red 1	Red 2
IFR1/IFC1	Puesto 2	Puesto 1
IFR2/IFC2	Puesto 1	Puesto 2
IFR3/IFC3	Puesto 1	Puesto 2
IFR4/IFC4	Puesto 2	Puesto 1
IFR5/IFC5	Puesto 1	Puesto 2

Esta clasificación puede ampliarse a un conjunto mayor de IFR o IFC, pues puede servir para una comparación detallada y para una comparación interna.

En algunos casos puede ser necesario obtener la puntuación estadística de red para cada servicio y para todos los servicios. En el Anexo A se presenta una técnica para hacerlo. Sin embargo, hay que tener en cuenta que este tipo de técnica sólo es válida si está plenamente descrita y se basa en supuestos y condiciones técnicamente sólidos.

Anexo A

Aplicación de la significación estadística al análisis comparativo de redes móviles

(Este anexo forma parte integrante de la presente Recomendación.)

Por análisis comparativo se entiende la comparación entre IFR o IFC que describen la calidad de funcionamiento de diversas redes de operador. Para ser significativa la comparación debe basarse en pruebas de significación estadística (pruebas hipotéticas) que dependen del tipo de IFR o IFC que se compara, continuos (por ejemplo, MOS, parámetros RF) y tasas (por ejemplo, tasa de completión o tasa de fallo).

En el primer caso, con la Ecuación A-1 se determina la diferencia significativa [UIT-T P.1401]:

$$Z = \text{StatDiff}/\sqrt{\text{std1}^2/\text{N1} + \text{std2}^2/\text{N2}} > Z_{th} \quad (\text{A-1})$$

siendo StatDiff la diferencia entre unidades comparadas, std1 y std2 sus desviaciones típicas y N1 y N2 el número total de muestras utilizadas para comparar cada unidad. Dicho de otro modo, si Z es superior a Z-ésima (sobre la base de una distribución gaussiana con más de 30 muestras y un nivel de confianza de NC%), StatDiff es una diferencia estadísticamente significativa con un nivel de confianza de NC%.

En el segundo caso, el tipo de relación IFR o IFC se describe mediante el número p de éxitos o fallos en relación con el número total de muestras. La diferencia significativa se obtiene con la Ecuación A-2 [b-UIT-T P.1401]

$$Z = \text{StatDiff}/\sqrt{p1*(1-p1)/\text{N1} + p2*(1-p2)/\text{N2}} > Z_{th} \quad (\text{A-2})$$

siendo p1 y p2 los números de éxitos o fallos de cada unidad comparada.

En el Cuadro A.1 se muestra la correspondencia entre los umbrales de significación, Z-ésima, con distintos niveles de confianza.

Cuadro A.1 – Correspondencia entre umbrales de significación con distintos niveles de confianza

NC%	90	95	96	97	98	99
Z-ésima	1,64	1,96	2,05	2,17	2,33	2,58

Si se dispone de menos de 30 muestras, se utilizará una distribución t-Student, en la que t-Student (n), siendo n = N-1 el número de grados de libertad y N el número total de muestras.

Cabe señalar que, además de la significación estadística, se han de utilizar los umbrales de diferencia específicos de IFR o IFC pertinentes cuando las diferencias son insignificantes o puedan estar dentro del margen de precisión de la medición de cada IFR o IFC. En otras Recomendaciones UIT-T de la serie sobre comparación de referencia se definen los umbrales de diferencia específicos de IFR o IFC pertinentes (THrelv).

Anexo B

Puntuación y clasificación estadísticas de la calidad de funcionamiento de una red

(Este anexo forma parte integrante de la presente Recomendación.)

En este anexo se describe el algoritmo para puntuar y clasificar la calidad de funcionamiento de extremo a extremo de las redes utilizado en los cálculos del Cuadro 3.

- Calcular los IFR (también pueden ser IFC) de extremo a extremo correspondientes al servicio considerado en cada red u operador:
 - IFR₁...IFR_i...IFR_N, i=1,n que pueden ser medias, medianas o proporciones (tasas).
- Crear una matriz comparativa para j=1,M redes (operadores) por servicio descrito por N IFR o IFC – véase el Cuadro B.1.

Cuadro B.1 – Matriz comparativa

	Red_1.....	Red_j.....	Red_M
IFR_1...	IFR _{1,1}	IFR _{1,j}	IFR _{1,M}
IFR _i ..	IFR _{i,1}	IFR _{i,j}	IFR _{i,M}
IFR _N	IFR _{N,1}	IFR _{N,j}	IFR _{N,M}

- Calcular la distancia de significación estadística para cada IFR_{i,j} de la matriz.
Para i=1,N
 - seleccionar el mejor valor, IFR_{i,best}, siendo "best" la j red con la mejor red
 - calcular la diferencia de significación estadística, StatDiff_{i,j}, entre cada IFR_{i,j} y el mejor valor, IFR_{i,best}, con las Ecuaciones B-1 o B-2, en función del tipo de elemento medido

$$\text{StatDiff}_{i,j} = \text{máx} \{0, (\text{KPI}_{i,\text{best}} - \text{KPI}_{i,j}) / \sqrt{(\text{std1}^2/N1 + \text{std2}^2/N2)} - Z_{th}\} \quad (\text{B-1})$$

$$\text{StatDiff}_{i,j} = \text{máx} \{0, (\text{KPI}_{i,\text{best}} - \text{KPI}_{i,j}) / \sqrt{p1*(1-p1)/N1 + p2*(1-p2)/N2} - Z_{th}\} \quad (\text{B-2})$$

NOTA – Z-ésima es F(0,05, N1, N2), el resultado estadístico con una significación del 95% y N1 y N2 grados de libertad.

Fin

- Determinar la red u operador con el puesto 1 ("mejor calidad de funcionamiento") para el servicio probado en la zona.

El puesto 1 se atribuye a la red con menor distancia estadísticamente significativa de calidad para el servicio probado de entre todas las redes u operadores considerados.

Para j=1,M

Puesto 1= Puesto(j) si StatDiffQuality_{mín}=mín(j=1,M) {SUM (i=1,N) {StatDiff_{i,j} * w_i}}

siendo StatDiff_{i,j} como se define en las Ecuaciones B-1 o B-2 y w_i las ponderaciones predefinidas para cada servicio, si se quieren utilizar ponderaciones. En caso contrario, pueden utilizarse ponderaciones unitarias iguales.

Fin

Determinar los puestos de todas las demás redes u operadores considerados en la campaña comparativa.

Para $j=1,M$

Si $\text{Dist}(j) = \max(0, \text{StatDiffQuality}/\text{StatDiffQuality}_{\text{mín}} - Z\text{-ésima}) = 0$

Puesto (j) = Puesto 1

(determina todas las redes con idéntico rendimiento estadístico al puesto 1)

O

Ordenar $\text{Dist}(j)$ de menor a mayor

Puesto (j) = posición en el vector $\text{Dist}(j)$

Fin

Cabe señalar que la clasificación debe efectuarse en función de la significación estadística de los umbrales de diferencia específicos de IFR o IFC pertinentes cuando las diferencias sean insignificantes o entren dentro del margen de precisión de la medición de cada IFR o IFC. En otras Recomendaciones UIT-T de la serie sobre comparación de referencia se definen los umbrales de diferencia específicos de IFR o IFC pertinentes (THrelv).

Apéndice I

Posible técnica para la puntuación y clasificación estadísticas de una red

(Este apéndice no forma parte integrante de la presente Recomendación.)

A veces puede ser conveniente determinar la puntuación estadística general de una red. Para ello se suele utilizar como criterio básico la calidad de funcionamiento de la red para cada tipo de servicio. Esa puntuación se ha de determinar en cada zona y acumulada en todas las zonas, como se muestra en la Figura 1.

La puntuación estadística por servicio puede definirse gracias a todos los IFR_i o IFC_i (i=1,N) de extremo a extremo considerados en relación con la calidad global de extremo a extremo del servicio analizado. Así, la puntuación puede definirse mediante la suma ponderada de la StatDiff_i (véase el Anexo B) de cada IFC en relación con los mejores resultados, como se describe en la cláusula 9.4. De ser posible, el valor StatDiff_i se corrige en término umbral de diferencia pertinente. El resultado final, StatScore, describe la calidad de funcionamiento de extremo a extremo de las redes consideradas en relación con la red con mejor calidad de funcionamiento.

$$\text{StatScore} = \Sigma(w_i * \text{StatDiff}_i)$$

Aquí, w_i es la ponderación de cada IFR o IFC relacionado con la calidad del servicio. Cuanto menor sea el StatScore, mejor será la calidad de funcionamiento (más cercana a la de la red con mejor calidad de funcionamiento) y más alto será el puesto correspondiente.

El Cuadro I.1 es una versión actualizada del Cuadro 3 en la que se han añadido algunas ponderaciones. Este ejemplo es meramente informativo, pues la definición de la ponderación queda fuera del alcance de esta Recomendación. Las ponderaciones también pueden ser unitarias, si se decide que todos los IFR o IFC tienen la misma importancia en el marco de la puntuación estadística general de la calidad de funcionamiento de la red. Sin embargo, cabe señalar que, aun cuando se utilizan ponderaciones unitarias iguales, el aumento o la reducción del número de IFR o IFC puede modificar la puntuación estadística de la red y arrojar resultados estadísticos distintos.

Por consiguiente, la puntuación y clasificación estadísticas de una red conforme a esta Recomendación sólo serán válidas si se cuenta con una descripción detallada y motivada de la selección y de las ponderaciones subyacentes de los IFR o IFC escogidas. Sin transparencia, la puntuación y clasificación estadísticas de la red no son válidas.

En el ejemplo del Cuadro I.1, la red 1 recibe el puesto 1, con una puntuación mínima de 2,15, de acuerdo con las ponderaciones aplicadas. Además, puede verse que en todos los casos las diferencias entre IFR o IFC de las dos redes son superiores al THrevl correspondiente, lo que supone que los resultados del análisis de significación estadística son válidos.

Cuadro I.1 – Ejemplo de puntuación y clasificación estadísticas

	Red 1				Red 2				
	IFR	Típica	N	StatDiff	IFR	Típica	N	StatDiff	THrevl
IFR1/IFC1	0,95	0,22	87	0,05	0,97	0,17	69	0,00	0,018
IFR2/IFC2	0,93	0,26	87	0,00	0,91	0,29	69	0,23	0,019
IFR3/IFC3	3,89	0,50	2 600	0,00	3,56	0,70	2 070	17,15	0,31
IFR4/IFC4	105,00	5,00	435	42,67	70,00	15,00	350	0,00	34
IFR5/IFC5	1 200,00	300,00	87	0,00	1 800,00	275,00	69	12,31	596
StatScore				2,15				5,83	
Puesto				1				2	

Una campaña comparativa completa puede estar destinada a calcular la puntuación estadística global de cada red considerando todos los servicios soportados ($j=1,M$ número de servicios soportados). Cabe señalar que, para ser conformes con esta Recomendación, toda puntuación global de ese tipo debe estar bien motivada y que todas las ponderaciones aplicadas a los servicios deben ser explícitamente transparentes.

La puntuación puede calcularse sumando todas las puntuaciones estadísticas de cada servicio, según los cálculos del Anexo B cuyo ejemplo se muestra en el Cuadro I.1. En ese caso, la red con mejor calidad de funcionamiento tendrá una puntuación estadística global de cero. Cuanto menor sea la puntuación, mejor será la calidad de funcionamiento de la red correspondiente, definida por la ponderación y suma de los IFR o IFC respectivos a cada servicio. La puntuación resultante de esa operación suele utilizarse para representar la clasificación global de las redes consideradas en relación con la de mejor calidad de funcionamiento:

$$\text{GlobalNetScore} = \Sigma(\text{Wser}_i * \text{StatScore}_j)$$

La definición de las ponderaciones Wserv_j para GlobalNetScore queda fuera del alcance de esta Recomendación, aunque se ofrecen algunas orientaciones.

En función de las políticas u objetivos comerciales del operador, GlobalNetScore puede calcularse como una suma ponderada de manera semejante a las ponderaciones utilizadas para el cálculo de StatScore . Las ponderaciones, así como el número o tipo de servicios considerados en el cálculo de GlobalNetScore dependerán de la elección o decisión de los operadores si lo que se hace es una comparación interna (Figura 1).

Otra posibilidad es decidir las ponderaciones en función del análisis estadístico de los usuarios (por ejemplo, financiación participativa), de los tipos de servicio utilizados o del porcentaje de utilización en distintas zonas.

Sin embargo, la validez de la puntuación global depende de la descripción y motivación detalladas de la selección de IFR o IFC y de las ponderaciones escogidas para cada servicio, así como del peso de cada uno de ellos en la nota global. Sin tal transparencia, la puntuación y clasificación estadísticas globales no serán válidas. Además, como ya se ha indicado en el cuerpo de esta Recomendación, las puntuaciones StatDiff subyacentes deben considerarse en función de la pertinencia de las diferencias entre servicios de los distintos IFR o IFC y debe tener el valor 0 si la diferencia puede ser significativa, pero no contribuye a mejorar la calidad de funcionamiento de la red desde la perspectiva del usuario.

Bibliografía

- [b-UIT-T E.802] Recomendación UIT -T E.802 (2007), *Marco y metodologías para la determinación y la aplicación de parámetros de calidad de servicio.*
- [b-UIT-T P.10] Recomendación UIT-T P.10/G.100 (2017), *Vocabulario sobre calidad de funcionamiento, calidad de servicio y calidad de la experiencia.*
- [b-UIT-T P.863] Recomendación UIT-T P.863 (2018), *Predicción de la calidad de escucha objetiva por percepción*
- [b-UIT-T P.1203] Recomendación UIT-T P.1203 (2017), *Evaluación paramétrica de la calidad basada en el tren de bits de los servicios audiovisuales de emisión de secuencias de descarga progresiva y adaptativa a través de un transporte fiable.*
- [b-UIT-T P.1401] Recomendación UIT-T P.1401 (2012), *Métodos, medición y procedimientos para la evaluación, calificación y comparación estadística de los modelos de predicción de la calidad objetiva.*
- [b-ETSI TR 102 581] ETSI TR 102 581, V1.2.1 (2015), *Speech processing, transmission and quality aspects (STQ); A study on the minimum additional required attenuation on the antenna path of the field test equipment.*
https://www.etsi.org/deliver/etsi_tr/102500_102599/102581/01.02.01_60/tr_102581v010201p.pdf
- [b-Mehta] Mehta, S. (2014). *Statistics topics*. CreateSpace. 160 pp.
- [b-Shaum] Spiegel, M.R., Schiller, J.J., Srinivasan, R.A. (2013). *Schaum's outlines: Probability and statistics*, 4th edition. New York, NY: McGraw-Hill. 424 pp.

SERIES DE RECOMENDACIONES DEL UIT-T

Serie A	Organización del trabajo del UIT-T
Serie D	Principios de tarificación y contabilidad y cuestiones económicas y políticas de las telecomunicaciones/TIC internacionales
Serie E	Explotación general de la red, servicio telefónico, explotación del servicio y factores humanos
Serie F	Servicios de telecomunicación no telefónicos
Serie G	Sistemas y medios de transmisión, sistemas y redes digitales
Serie H	Sistemas audiovisuales y multimedia
Serie I	Red digital de servicios integrados
Serie J	Redes de cable y transmisión de programas radiofónicos y televisivos, y de otras señales multimedia
Serie K	Protección contra las interferencias
Serie L	Medio ambiente y TIC, cambio climático, ciberdesechos, eficiencia energética, construcción, instalación y protección de los cables y demás elementos de planta exterior
Serie M	Gestión de las telecomunicaciones, incluida la RGT y el mantenimiento de redes
Serie N	Mantenimiento: circuitos internacionales para transmisiones radiofónicas y de televisión
Serie O	Especificaciones de los aparatos de medida
Serie P	Calidad de la transmisión telefónica, instalaciones telefónicas y redes de líneas locales
Serie Q	Conmutación y señalización, y mediciones y pruebas asociadas
Serie R	Transmisión telegráfica
Serie S	Equipos terminales para servicios de telegrafía
Serie T	Terminales para servicios de telemática
Serie U	Conmutación telegráfica
Serie V	Comunicación de datos por la red telefónica
Serie X	Redes de datos, comunicaciones de sistemas abiertos y seguridad
Serie Y	Infraestructura mundial de la información, aspectos del protocolo Internet, redes de próxima generación, Internet de las cosas y ciudades inteligentes
Serie Z	Lenguajes y aspectos generales de soporte lógico para sistemas de telecomunicación