

ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA DE TELECOMUNICACIÓN
UNIVERSIDAD POLITÉCNICA DE CARTAGENA



Proyecto Fin de Carrera

Encaminamiento dinámico en redes de conmutación de circuitos



AUTOR: Juan Pedro Muñoz Gea
DIRECTOR: Josemaría Malgosa Sanahuja

Septiembre / 2.005



Autor	Juan Pedro Muñoz Gea
E-mail del Autor	juanpemg@upct.universia.es
Director	Josemaría Malgosa Sanahuja
E-mail del Director	josem.malgosa@upct.es
Título del PFC	Encaminamiento dinámico en redes de conmutación de circuitos
Descriptorios	Redes de conmutación de circuitos, encaminamiento dinámico, inestabilidad, reserva de troncales, encaminamiento aleatorio, encaminamiento por ruta menos cargada
<p>Resumen</p> <p>Uno de los principales aspectos a tener presente en el diseño de las redes de conmutación de circuitos (redes de voz) son los algoritmos y protocolos de señalización; especialmente aquellos cuyo propósito consiste en establecer el circuito entre el llamante y el llamado. En este sentido este proyecto trata de estudiar diferentes algoritmos dinámicos para el establecimiento de la llamada, con objeto de reducir al máximo la probabilidad de pérdida (o bloqueo) que percibe el usuario del servicio.</p> <p>En toda red con encaminamiento dinámico no jerárquico aparecen márgenes de inestabilidad, es decir, rangos de tráfico ofrecido para los cuales existe más de una solución en la probabilidad de pérdida. La reserva de un número pequeño de troncales para uso exclusivo del tráfico directo entre cada par de nodos soluciona el problema de la inestabilidad. Se demostrará la existencia de las inestabilidades y su eliminación mediante la reserva de troncales de forma teórica, y se comprobará posteriormente mediante simulación, a través de herramientas desarrolladas en lenguaje C.</p> <p>Se analizará la clasificación de los distintos métodos de encaminamiento dinámico, y se evaluarán las prestaciones de un algoritmo de encaminamiento aleatorio, como es el DAR, y de un algoritmo de encaminamiento por ruta menos cargada, como es el LLR. En la medida de lo posible, se extrapolarán los resultados obtenidos para aplicarlos en redes del tipo VoIP.</p>	
Titulación	Ingeniería de Telecomunicación
Intensificación	Planificación y Gestión de Telecomunicaciones
Departamento	Tecnologías de la Información y las Comunicaciones
Fecha de Presentación	Septiembre - 2.005

Tabla de Contenidos

CAPÍTULO 1..... 1

1.1. INTRODUCCIÓN	1
1.2. ENCAMINAMIENTO JERÁRQUICO	2
1.3. ENCAMINAMIENTO NO JERÁRQUICO	3
1.4. ENCAMINAMIENTO DIRECTO.....	5
1.5. NOTACIÓN.....	6
1.6. SEÑALIZACIÓN DE CONTROL.....	8
1.6.1. FUNCIONES DE SEÑALIZACIÓN	8
1.6.2. SEÑALIZACIÓN POR CANAL COMÚN.....	9
1.6.3. EL PROTOCOLO SS7	10
1.7. OBJETIVOS.....	11

CAPÍTULO 2..... 13

2.1. INTRODUCCIÓN	13
2.2. MODELO EFP	13
2.2.1. MODELADO GENERAL.....	13
2.2.2. PARTICULARIZACIÓN A UNA RED SIMÉTRICA TOTALMENTE CONECTADA	15
2.3. RESOLUCIÓN ITERATIVA DE LA ECUACIÓN DE RED	16
2.4. RESULTADOS TEÓRICOS	18
2.4.1. RED ESTABLE.....	18
2.4.2. DESESTABILIZACIÓN POR NÚMERO DE RUTAS ALTERNATIVAS	20
2.4.3. DESESTABILIZACIÓN POR CAPACIDAD EN LOS ENLACES	22
2.4.4. NÚMERO MEDIO DE ENLACES OCUPADOS	22
2.5. RESULTADOS DE SIMULACIÓN.....	24
2.6. MÉTODO DE RESERVA DE TRONCALES	26
2.6.1. DESARROLLO TEÓRICO	26
2.6.2. RESOLUCIÓN ITERATIVA.....	29
2.6.3. RESULTADOS	30
2.7. MÉTODOS DE ENCAMINAMIENTO DINÁMICO	32
2.7.1. BREVE HISTORIA DEL ENCAMINAMIENTO DINÁMICO	32
2.7.2. CLASIFICACIÓN DE LOS MÉTODOS DE ENCAMINAMIENTO ALTERNATIVO	33
2.7.3. IMPLEMENTACIONES ACTUALES DE ENCAMINAMIENTO DINÁMICO	36
2.7.4. OTRAS CLASIFICACIONES	36

CAPÍTULO 3..... 39

3.1. INTRODUCCIÓN	39
3.2. DESCRIPCIÓN DEL DAR (DYNAMIC ALTERNATIVE ROUTING)	39
3.3. PRESTACIONES EN REDES MULTICLASE	40
3.3.1. DISTRIBUCIÓN DE LOS TIEMPOS DE SERVICIO Y ENTRE LLAMADAS	40
3.4. COMPARACIÓN DE PRESTACIONES CON EL ENCAMINAMIENTO DIRECTO.....	41
3.4.1. RESULTADOS GENERALES.....	41
3.4.2. INFLUENCIA DEL TAMAÑO DE LA RED	44
3.5. OTROS ALGORITMOS ALEATORIOS	45

CAPÍTULO 4	47
4.1. INTRODUCCIÓN	47
4.2. DESCRIPCIÓN DEL LLR (LEAST LOADED ROUTING)	47
4.2.1. DESCRIPCIÓN GENERAL	47
4.2.2. IMPLEMENTABILIDAD DEL LLR.....	48
4.3. PRESTACIONES EN REDES MULTICLASE	48
4.3.1. DISTRIBUCIÓN DE LOS TIEMPOS DE SERVICIO Y ENTRE LLAMADAS	48
4.4. COMPARACIÓN DE PRESTACIONES CON EL ENCAMINAMIENTO DIRECTO	49
4.4.1. RESULTADOS GENERALES.....	49
4.4.2. INFLUENCIA DEL TAMAÑO DE LA RED	51
4.5. ANÁLISIS DEL RE-ENCAMINAMIENTO EN LLR	52
4.6. OTROS ALGORITMOS DE RUTA MENOS CARGADA	53
4.6.1. APLICACIÓN DE LAS REDES NEURONALES AL ENCAMINAMIENTO DINÁMICO EN REDES DE CONMUTACIÓN DE CIRCUITOS	53
 CAPÍTULO 5	 55
5.1. INTRODUCCIÓN	55
5.2. COMPARACIÓN ENTRE LOS TIPOS DE ENCAMINAMIENTO	55
5.3. COMPARACIÓN DAR-LLR	58
5.3.1. COMPARACIÓN DESCRIPTIVA	58
5.3.2. COMPARACIÓN DE PRESTACIONES CON EL ENCAMINAMIENTO DIRECTO	59
5.3.3. INFLUENCIA DEL TAMAÑO DE LA RED	62
5.3.4. DISTRIBUCIÓN DE LOS TIEMPOS DE SERVICIO Y ENTRE LLAMADAS	63
5.4. CONCLUSIONES	64
 CAPÍTULO 6	 65
6.1. INTRODUCCIÓN	65
6.2. REDES ATM	65
6.3. REDES IP	66
6.4. MPLS	67
6.5. SIMULACIÓN	69
6.5.1. RESULTADOS DE SIMULACIÓN	72
6.6. CONCLUSIONES	77
 CAPÍTULO 7	 79
7.1. CONCLUSIONES	79
7.2. LÍNEAS FUTURAS	80
 ANEXO	 81
A.1. INTRODUCCIÓN	81
A.2. MÉTODO DE LA TRANSFORMACIÓN INVERSA	81
A.3. DISTRIBUCIÓN DETERMINISTA	82
A.4. DISTRIBUCIÓN UNIFORME	82
A.5. DISTRIBUCIÓN EXPONENCIAL	83
A.6. DISTRIBUCIÓN TRIANGULAR	84
A.7. DISTRIBUCIÓN DE RAYLEIGH	87

A.8. DISTRIBUCIÓN NORMAL	88
<u>ACRÓNIMOS.....</u>	<u>91</u>
<u>BIBLIOGRAFÍA Y REFERENCIAS</u>	<u>93</u>

Índice de Figuras

Figura 1.1 – Encaminamiento jerárquico.	2
Figura 1.2 – Red no jerárquica.	3
Figura 1.3 – Red backbone no jerárquica.	3
Figura 1.4 – Posibles rutas entre un par de nodos con rutas secundarias de dos enlaces máximo.	4
Figura 1.5 – Encaminamiento directo.	5
Figura 1.6 – Modelado del encaminamiento directo.	5
Figura 1.7 – Comportamiento Erlang-B.	7
Figura 2.1 – Aspecto de $f(p)$ para diferentes valores de tráfico ofrecido. Red simétrica, totalmente conectada, $N=5$, $M=3$, $c=100$	16
Figura 2.2 – Programa MATLAB para la resolución de la ecuación de red.	17
Figura 2.3 – Comportamiento de una red simétrica totalmente conectada ($N=5$, $c=10$) frente al tráfico normalizado ofrecido entre cada par de nodos para diferente número de rutas alternativas (M). (a) Tráfico normalizado ofrecido a cada enlace. (b) Tráfico normalizado cursado por cada enlace. (c) Tráfico normalizado cursado por la red entre cada par de nodos. (d) Probabilidad de bloqueo de un enlace. (e) Probabilidad de bloqueo de ruta.	18
Figura 2.4 – Ampliación de la figura 2.3c.	19
Figura 2.5 – Ampliación de la figura 2.3e.	20
Figura 2.6 – Comportamiento de una red simétrica totalmente conectada ($N=5$, $c=100$) frente al tráfico normalizado ofrecido entre cada par de nodos para diferente número de rutas alternativas. (a) Tráfico normalizado ofrecido a cada enlace. (b) Tráfico normalizado cursado por cada enlace. (c) Tráfico normalizado cursado por la red entre cada par de nodos. (d) Probabilidad de bloqueo de un enlace. (e) Probabilidad de bloqueo de ruta.	21
Figura 2.7 – Comportamiento de una red simétrica totalmente conectada ($M=1$) frente al tráfico normalizado ofrecido entre cada par de nodos para enlaces de diferente capacidad. (a) Tráfico normalizado ofrecido a cada enlace. (b) Tráfico normalizado cursado por cada enlace. (c) Tráfico normalizado cursado por la red entre cada par de nodos. (d) Probabilidad de bloqueo de un enlace. (e) Probabilidad de bloqueo de ruta.	23
Figura 2.8 – Número medio de enlaces ocupados en una red simétrica totalmente conectada ($N=5$, $c=100$) frente al tráfico normalizado ofrecido entre cada par de nodos para diferente número de rutas alternativas.	24
Figura 2.9 – Programa AWK para procesamiento de resultados.	25
Figura 2.10 – Evolución temporal del GoS de una red simétrica totalmente conectada ($N=5$, $c=10$, $M=3$).	25
Figura 2.11 – Evolución temporal del GoS de una red simétrica totalmente conectada ($N=5$, $c=100$, $M=3$).	26
Figura 2.12 – Diagrama de estados de un enlace con reserva de troncales.	26
Figura 2.13 – Comportamiento de una red simétrica totalmente conectada ($N=5$, $c=100$, $M=3$) frente al tráfico normalizado ofrecido entre cada par de nodos para diferentes valores de reserva. (a) Tráfico normalizado ofrecido a cada enlace. (b) Tráfico normalizado cursado por cada enlace. (c) Tráfico normalizado cursado por la red entre cada par de nodos. (d) Probabilidad de bloqueo de un enlace. (e) Probabilidad de bloqueo de ruta.	31
Figura 2.14 – Número medio de enlaces ocupados en una red simétrica totalmente conectada ($N=5$, $c=100$, $M=3$) frente al tráfico normalizado ofrecido entre cada par de nodos para diferente número de rutas reservadas.	30
Figura 2.15 – Implementaciones de métodos de encaminamiento dinámico.	36
_Toc112493003	
Figura 3.1 – Modo de operación del DAR.	39

Figura 3.2 – Dependencia del GoS con la distribución del tráfico en una red totalmente conectada simétrica; $N=5$, $c=10$. (a) GoS para tiempo entre llamadas exponencial y tiempo de servicio variable. (b) GoS para tiempo de servicio exponencial y tiempo entre llamadas variable.	41
Figura 3.3 – Prestaciones del encaminamiento directo y del DAR en alta carga.	42
Figura 3.4 – Prestaciones de DAR en relación al encaminamiento directo. (a) GoS de la red para diferentes valores medios de tráfico ofrecido normalizado entre cada par de nodos. (b) Tráfico cursado para diferentes GoSs. (c) Ganancia absoluta en GoS. (d) Ganancia absoluta en tráfico cursado. (e) Ganancia relativa en GoS. (f) Ganancia relativa en tráfico cursado.	43
Figura 3.5 – GoS del DAR para diferentes tamaños de red.	45
Figura 3.6 – GoS del DAR para diferentes tamaños de red.	45
Figura 4.1 – LLR: todos los nodos son posibles nodos intermedios.	47
Figura 4.2 – Dependencia del GoS con la distribución del tráfico en una red totalmente conectada simétrica; $N=5$, $c=10$. (a) GoS para tiempo entre llamadas exponencial y tiempo de servicio variable. (b) GoS para tiempo de servicio exponencial y tiempo entre llamadas variable.	48
Figura 4.3 – Prestaciones del encaminamiento directo y del LLR en alta carga.	49
Figura 4.4 – Prestaciones de LLR en relación al encaminamiento directo. (a) GoS de la red para diferentes valores medios de tráfico ofrecido normalizado entre cada par de nodos. (b) Tráfico cursado para diferentes GoSs. (c) Ganancia absoluta en GoS. (d) Ganancia absoluta en tráfico cursado. (e) Ganancia relativa en GoS. (f) Ganancia relativa en tráfico cursado.	50
Figura 4.5 – GoS del LLR para diferentes tamaños de red.	51
Figura 4.6 – GoS del LLR para diferentes tamaños de red.	52
Figura 4.7 – Camino alternativo entre un par de nodos.	52
Figura 5.1 – Prestaciones del encaminamiento directo, del DAR y del LLR en alta carga.	59
Figura 5.2 – GoS de la red para diferentes valores medios de tráfico ofrecido medio normalizado.	60
Figura 5.3 – Ganancia absoluta en GoS de la red para diferentes valores medios de tráfico ofrecido medio normalizado.	60
Figura 5.4 – Ganancia relativa en GoS de la red para diferentes valores medios de tráfico ofrecido medio normalizado.	61
Figura 5.5 – Tráficos cursados para diferentes GoSs.	61
Figura 5.6 – Ganancia absoluta en tráfico cursado.	61
Figura 5.7 – Ganancia relativa en tráfico cursado.	62
Figura 5.8 – GoS del DAR para diferentes tamaños de red.	62
Figura 5.9 – GoS del LLR para diferentes tamaños de red.	63
Figura 6.1 – Protocolos implicados en la transmisión de VoIP.	69
Figura 6.2 – Red simétrica totalmente conectada de cinco nodos.	70
Figura 6.3 – Modelización de las colas para controlar el retraso inicial y el <i>jitter</i>	71
Figura 6.4 – Prestaciones del encaminamiento directo y del dinámico, para dos valores de K	72
Figura 6.5 – GoS de la red para diferentes valores medios de tráfico ofrecido entre cada par de nodos. ...	73
Figura 6.6 – GoS de la red para diferentes valores medios de tráfico ofrecido entre cada par de nodos. ...	74
Figura 6.7 – Ganancia absoluta en GoS.	74
Figura 6.8 – Ganancia relativa en GoS.	74
Figura 6.9 – Utilización del encaminamiento directo y del dinámico, para dos valores de K	75
Figura 6.10 – Utilización de la red para diferentes valores medios de tráfico ofrecido entre cada par de nodos.	76
Figura A.1 – Ejemplo: FDP y fdp de Rayleigh.	82
Figura A.2 – FDP y fdp exponenciales con $m=1$	83
Figura A.3 – Función de distribución triangular.	84
Figura A.4 – Separación de la función de distribución triangular.	85
Figura A.5 – FDP y fdp triangulares de $a = 1$, $c = 2$ y $b = 4$	87
Figura A.6 – FDP y fdp de Rayleigh con $\sigma = 1$	88
Figura A.7 – FDP y fdp normales o gaussianas de $m = 0$ y $\sigma = 1$	88

Índice de Tablas

Tabla 2.1 – Raíces de la figura 2.1.....	16
Tabla 2.2 – Estado de las redes de encaminamiento dinámico.	38
Tabla 3.1 – Ganancias absoluta y relativa máximas para GoS y para tráfico.	44
Tabla 3.2 – Tráficos y ganancias absoluta y relativa para GoS máximo.....	44
Tabla 3.3 – Número de circuitos reservados para tráfico directo, y tráfico ofrecido para alcanzar GoS=2% con encaminamiento directo.	44
Tabla 4.1 – Ganancias absoluta y relativa máximas para GoS y para tráfico.	51
Tabla 4.2 – Tráficos y ganancias absoluta y relativa para GoS máximo.....	51
Tabla 6.1 – Ganancias absoluta y relativa máximas para GoS.....	75
Tabla 6.2 – Tráficos ofrecidos y utilizaciones para GoS máximo.....	76

Capítulo 1

Introducción al encaminamiento en redes de pérdidas

1.1. Introducción

Cuando se establece una llamada, la red debe de encontrar una ruta desde el abonado que llama al llamado, que pasa a través de varios conmutadores y enlaces. Existen dos requisitos fundamentales para la arquitectura de red: la eficiencia y la flexibilidad. Hay que minimizar la cantidad de conmutadores y enlaces en la red, teniendo en cuenta que debe de ser capaz de aceptar toda la carga esperada, o sea, la carga promedio en horas punta a lo largo del día. Desde un punto de vista práctico es necesario ser capaz de gestionar esta cantidad de tráfico. Desde el punto de vista de costes sería deseable poder gestionar esta carga con el menor equipamiento posible.

Otro requisito es la flexibilidad para adaptarse a las variaciones de tráfico con un nivel de servicio razonable, incluso en horas punta o en condiciones excepcionales. La estrategia de encaminamiento debe ser, por lo tanto, un compromiso entre eficiencia y flexibilidad.

Antiguamente, los conmutadores de una red se organizaban en una estructura en árbol o *jerárquica*. La ruta se establecía a través del árbol, comenzando en el abonado que llamaba, llegando hasta el primer nodo común y después hasta el abonado llamado. Para proporcionar flexibilidad a la red se incluían en el árbol enlaces de alta capacidad adicionales para conectar entre sí centrales con altos volúmenes de tráfico. Esta aproximación es estática. La incorporación de enlaces de alta capacidad proporciona redundancia y capacidad extra, pero no es capaz de adaptarse a condiciones cambiantes e incluso a altas cargas de trabajo.

En la actualidad los operadores han pasado de una aproximación jerárquica a una *no jerárquica* en donde las decisiones de encaminamiento dependen en cada instante del tráfico actual. Esta arquitectura es más compleja, pero más flexible porque hay más rutas alternativas.

Un ejemplo de encaminamiento en una red *no jerárquica* es el *encaminamiento alternativo*, en donde cada conmutador dispone de un conjunto de rutas prefijadas en orden de preferencia para cada destino. Si existe una conexión directa, ésta es la elegida. Si no está disponible esta línea se prueba con la segunda alternativa y así sucesivamente. Las secuencias de encaminamiento –conjunto de rutas intentadas– reflejan un análisis basado en patrones de tráfico conocidos y se diseñan para optimizar la utilización de los recursos de la red, con el objetivo de disminuir la probabilidad de pérdida –o bloqueo– percibida por el usuario. Este parámetro también recibe el nombre de *grado de servicio* (GoS).

Si sólo se define una secuencia de encaminamiento para cada pareja origen–destino tenemos un *encaminamiento alternativo fijo* (o *encaminamiento fijo*), aunque es más frecuente el *encaminamiento alternativo dinámico* (*encaminamiento dinámico* o *encaminamiento dinámico no jerárquico*), en donde se utiliza un conjunto diferente de rutas en instantes distintos de tiempo, con objeto de aprovechar las distintas condiciones de tráfico en las diferentes zonas horarias y en los distintos periodos del día. Por tanto, la decisión de encaminamiento se basa tanto en el estado del tráfico actual –una ruta se descartará si está ocupada– como en patrones de tráfico conocidos –que determinan la secuencia de rutas a considerar.

1.2. Encaminamiento jerárquico

El encaminamiento mediante centrales jerarquizadas fue introducido por AT&T en 1930. En esos tiempos, operadores especializados realizaban las conexiones adecuadas de forma manual (no fue hasta los años 50 cuando el encaminamiento se comenzó a realizar de forma automática). Con pocas variaciones, este método de encaminamiento se mantiene activo en la actualidad en numerosos países, aunque a finales de los 80 e inicios de los 90 comenzó a ser reemplazado por el encaminamiento no jerárquico.

Los sistemas basados en encaminamiento jerárquico dividían recursivamente el territorio al que se pretendía dar cobertura en regiones, definiendo así diversas jerarquías a las que hay asociadas centrales de conmutación de una determinada clase. En la figura 1.1 se muestra un ejemplo en el que se aprecian cinco jerarquías, así como las conexiones entre las mismas. Las llamadas entre dos centrales locales se intentan cursar por la *troncal directa* que las une, la línea de trazo grueso en la figura. Cuando la *troncal directa* está ocupada, la llamada se desvía hacia una de las *centrales intermedias*, usando una de las *troncales alternativas* (con línea de puntos) y seguidamente una *troncal final* (línea de trazo fino), que une una central con su inmediata inferior. Si esta *troncal alternativa* tampoco puede cursar la llamada, se intenta por la otra central intermedia y posteriormente, si fuera necesario, por una troncal entre las dos centrales intermedias. Sucesivos bloqueos hacen que el mismo proceso se repita, cada vez usando recursos de más alta jerarquía. Cuando en un determinado instante del proceso de encaminamiento una llamada no puede ser servida por un nivel jerárquico, y todas las *troncales finales* (línea de trazo fino) de conexión con el nivel superior están ocupadas, la llamada no puede ser atendida y se rechaza.

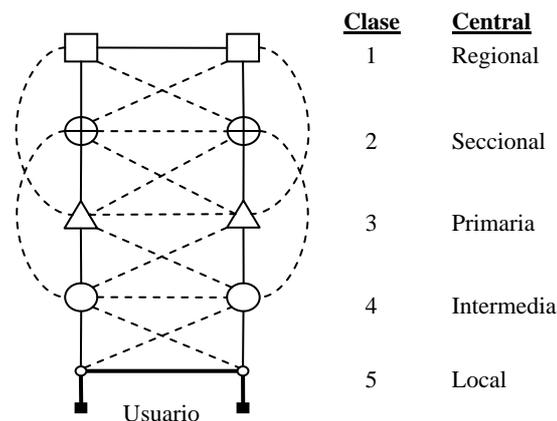


Figura 1.1 – Encaminamiento jerárquico.

Una de las ventajas que presenta el encaminamiento jerárquico es que es posible implementarlo con la tecnología electrónica tradicional, por lo que no es necesaria la utilización de los nuevos sistemas de conmutadores con programas de control.

Sin embargo, el encaminamiento jerárquico tiene una serie de limitaciones:

- Se requieren uno o varios conmutadores intermedios con funciones de tratar tráfico en tránsito.
- El establecimiento del circuito, con los mensajes de señalización correspondientes, es muy laborioso –pero todavía sin grandes problemas porque la búsqueda del camino sigue un esquema fijo determinado completamente por la jerarquía–. Una conexión completa puede involucrar un gran número de dispositivos intermedios, usando muchos recursos de la red para una única conexión.
- El esquema de numeración debe ser acorde con la estructura jerárquica.
- Incluso aunque la topología de la red pueda permitir un gran número de alternativas de encaminamiento, las restricciones sobre cómo se deben seleccionar las rutas limita la elección actual a un número relativamente pequeño.

- Los patrones de encaminamiento no se pueden cambiar de acuerdo a la hora del día o a los patrones de tráfico de la red.

Todas estas deficiencias están motivadas por el uso de una lógica de decisión relativamente simple en cada nodo, sin disponer de un conocimiento completo del estado de la red. La introducción de los conmutadores con programas de control almacenados y la señalización por canal común cambió la situación anterior, ya que hicieron posible procesos de encaminamiento más eficientes.

1.3. Encaminamiento no jerárquico

Como se ha visto en el punto anterior, el encaminamiento jerárquico tiene carácter estático. En la época de las redes jerárquicas, los ordenadores no se habían desarrollado y las reglas de encaminamiento debían ser lo suficientemente sencillas como para que fueran llevadas a cabo manualmente. La principal causa que ha motivado la migración hacia redes no jerárquicas, es decir, redes en las que todas las centrales de conmutación pertenecen al mismo nivel, ha sido el desarrollo tecnológico. El posterior y rápido desarrollo de las redes con programas de control permitió abandonar la rigidez estática de las redes jerárquicas y avanzar hacia el *encaminamiento dinámico no jerárquico*.

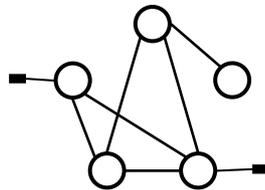


Figura 1.2 – Red no jerárquica.

En las redes no jerárquicas, los conmutadores *backbone* sustituyen a las centrales de conmutación jerárquicas. Todos estos conmutadores tienen la misma responsabilidad a la hora de encaminar las llamadas, ya que disponen de enlaces lógicos directos entre cada par de conmutadores *backbone* de la red. En este caso, el encaminamiento se realiza de la forma siguiente: si el origen y el destino de la llamada se encuentran dentro de la misma central local, la conexión se realiza de forma directa. Si la llamada se realiza entre centrales locales conectadas directamente, se utiliza el camino entre las dos centrales locales. En cualquier otro caso, la llamada se debe encaminar hacia la red *backbone*; en esta situación deben tomarse decisiones de encaminamiento más importantes, teniendo en cuenta que todos los conmutadores pertenecen al mismo nivel y que es posible cualquier tipo de conexión entre ellos.

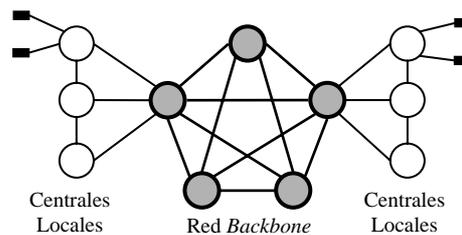


Figura 1.3 – Red backbone no jerárquica.

El encaminamiento no jerárquico –o encaminamiento alternativo– en redes *backbone* logra una mejora notable de las prestaciones, adaptándose a las variaciones del tráfico debidas, entre otras causas, a variación de la carga en determinadas horas del día, en fines de semana, días determinados del año, estaciones, etc. Otro motivo para esta mejora se produce en redes amplias, y que por lo tanto abarcan diferentes franjas horarias simultáneamente, ya que se produce un efecto de no coincidencia de la hora cargada, por lo que una parte de la red puede estar saturada mientras que otra dispone de recursos libres, que el encaminamiento dinámico debe aprovechar, cursando llamadas de la zona saturada a través de centrales de la zona libre.

Otros factores que también han marcado la evolución de estos nuevos métodos han sido:

- La *mejora de las prestaciones* de la red –*grado de servicio*–, a medida que la red puede adaptarse dinámicamente a variaciones de la carga y fallos de alguno de sus elementos.
- La introducción de *nuevos servicios*, gracias a la adaptación dinámica del encaminamiento a diferentes patrones de tráfico.
- La *evolución tecnológica* en transmisión, conmutación y gestión de la red, redundando en redes más eficientes y automatizadas.

Los algoritmos que serán objeto de estudio en los capítulos posteriores se engloban dentro del *encaminamiento dinámico no jerárquico*, aunque se basen en filosofías diferentes. Así,

- El DAR intenta cursar las llamadas entre cada par de nodos origen-destino por la ruta directa; si ésta está bloqueada, lo intenta por la ruta alternativa asociada al par de nodos considerados; si ésta también está bloqueada la llamada se pierde, y una nueva ruta alternativa se elige de forma aleatoria entre todas las posibles.
- El LLR también intenta la ruta directa como primera opción; como segunda opción, se intenta aquella ruta alternativa en la que el máximo de la ocupación de sus dos enlaces es menor; una llamada sólo será rechazada cuando no hay ninguna ruta que pueda cursarla.

Se asumirá que cualquier ruta está formada por uno (ruta directa) o dos enlaces. El hecho de limitar la ruta alternativa a *dos saltos* tiene dos motivos:

- Uno relacionado con la implementación del algoritmo, ya que resulta mucho más sencilla.
- Otro basado en la escasa mejora de prestaciones que se obtiene permitiendo rutas de tres enlaces, debido a que una sola llamada ocupe varios enlaces. Se comprobará, incluso, que cuando la red está funcionando por encima de un cierto nivel de carga, el hecho de permitir rutas alternativas distintas a la directa provoca un empeoramiento en el grado de servicio.

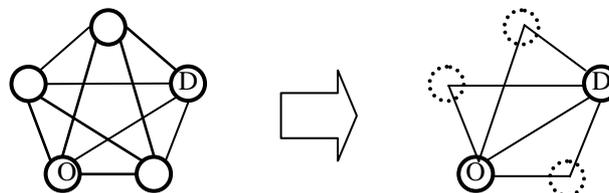


Figura 1.4 – Posibles rutas entre un par de nodos con rutas secundarias de dos enlaces máximo.

El encaminamiento no jerárquico presenta una serie de ventajas:

- En caso de un incremento imprevisto de tráfico, esta sobrecarga se distribuye por todos los enlaces equilibrando la situación.
- En una red con un número suficientemente grande de nodos, los incrementos imprevistos de tráfico casi no causan ningún incremento en el GoS predefinido.
- La red es más robusta contra averías.

Pero también presenta una serie de limitaciones:

- El sistema de señalización es más complicado.
- Se pueden presentar inestabilidades y oscilaciones en la distribución de tráfico.

1.4. Encaminamiento directo

En este apartado se aborda el estudio del encaminamiento que conecta los dos extremos remotos que desean establecer la comunicación mediante el enlace directo que los une. Con esta premisa, aunque estrictamente quizá se puede seguir hablando de encaminamiento, no es aplicable el adjetivo de dinámico ya que es totalmente estático. Toda llamada entre un determinado par de nodos va a intentar cursarse sólo por la ruta directa.

Sin embargo, resulta de gran interés un análisis de las prestaciones de este método de encaminamiento. En primer lugar, porque, dado que constituye la forma más sencilla de encaminar, en este proyecto se tomará como referencia a la hora de comparar las prestaciones de los algoritmos dinámicos que también serán objeto de estudio. En segundo lugar, se verá que gran parte de los algoritmos que se han estudiado, incluyendo algunos implementados en redes existentes, presentan problemas sobre un cierto nivel de carga, a partir del cual sólo se emplea el encaminamiento directo.

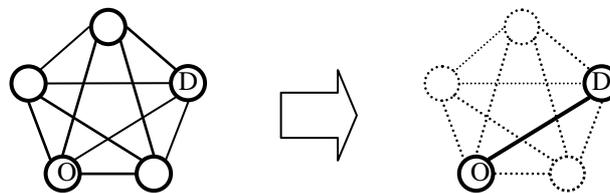


Figura 1.5 – Encaminamiento directo.

Dado que el encaminamiento directo sólo emplea las troncales que unen directamente los nodos que se han de conectar, y que el tráfico entre ellos es uno de los parámetros de la simulación, puede abordarse el estudio teórico de este encaminamiento como un proceso de nacimiento y muerte de tantos estados como troncales existen en el enlace, tal y como se muestra en la figura 1.6. Se asume un proceso de Poisson en el que los tiempos entre llegadas y de duración del servicio se distribuyen exponencialmente con medias $1/\lambda$ y $1/\mu$, respectivamente. El hecho de asumir que las tasas de llegada son iguales independientemente del estado del enlace implica la existencia de una población de usuarios infinita o muy elevada, suposición aceptable dentro del rango de redes que se va a estudiar, permitiendo modelar el sistema como un cola M/M/C/C en lugar de emplear modelos binomiales de población finita.

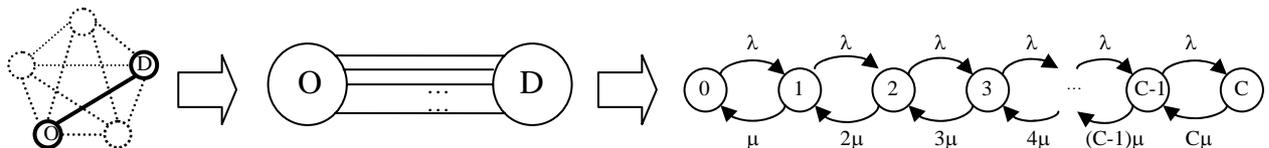


Figura 1.6 – Modelado del encaminamiento directo.

Visto desde el nivel de red, un enlace puede ser considerado como un conjunto de troncales con el mismo origen y destino. En un principio, estas troncales estaban constituidas por circuitos físicos eléctricamente independientes. Puede seguir manejándose este concepto, ya que da una idea más intuitiva del enlace, aunque actualmente las troncales se componen de ranuras temporales en tramas TDM, franjas espectrales en tramas FDM o grupos de celdas en conmutación ATM.

En cualquier caso, el estudio puede hacerse planteando C estados posibles, siendo C la cantidad de circuitos del enlace. Se está suponiendo la existencia de un solo servicio, por lo que el número de circuitos viene dado por la relación entre la capacidad del enlace y la velocidad del servicio. El desarrollo teórico conduce a que la probabilidad de estar en el estado n es

$$P(n) = \frac{(\lambda / \mu)^n / n!}{\sum_{i=0}^C (\lambda / \mu)^i / i!} \quad (1.1)$$

En virtud del principio PASTA (*Poisson Arrivals See Time Averages*), ya que el proceso de llegadas es de Poisson, la probabilidad de pérdida coincide con la probabilidad de bloqueo, es decir, con la probabilidad de que todas las troncales estén ocupadas. Por lo tanto

$$P_l = P(C) = \frac{(\lambda / \mu)^C / C!}{\sum_{i=0}^C (\lambda / \mu)^i / i!} \quad (1.2)$$

que es conocida como la *fórmula Erlang-B*, en honor al matemático danés que la propuso; en adelante, se utilizará la notación

$$P_l = B(A, C) \quad (1.3)$$

siendo $A = \lambda / \mu$ el tráfico total ofrecido al enlace. En las gráficas de la figura 1.7 se muestra el comportamiento de la Erlang-B mediante las seis relaciones que unen a las tres variables que intervienen. Puede comprobarse que, para una probabilidad de pérdida fijada, mayor tráfico requiere un mayor número de circuitos (figura a) y viceversa; para un tráfico fijado, el número de circuitos y la probabilidad de pérdida son inversamente proporcionales (figura c), etc.

Dos de las ventajas del encaminamiento directo son:

- No se requiere ningún conmutador intermedio con funciones de tránsito.
- El establecimiento del circuito, con los mensajes de señalización correspondientes, es muy sencillo.

De la misma forma, dos inconvenientes del encaminamiento directo son:

- Para valores de tráfico muy bajos y una probabilidad de pérdida muy restringida, el encaminamiento directo lleva a un bajo grado de uso de los circuitos.
- En caso de una avería del enlace los dos nodos se desconectan.

Algunas de las aplicaciones del encaminamiento directo son:

- En redes privadas y corporativas.
- En redes públicas:
 - Como parte del encaminamiento jerárquico, en la parte de interconexión de nodos del nivel superior con una probabilidad de pérdida muy limitada baja.
 - En la parte baja, entre nodos del nivel inferior con una probabilidad de pérdida (desbordamiento) alta.

1.5. Notación

En este apartado se presenta una parte de la notación matemática que será utilizada en el resto del proyecto. Se define el conjunto de nodos de red N

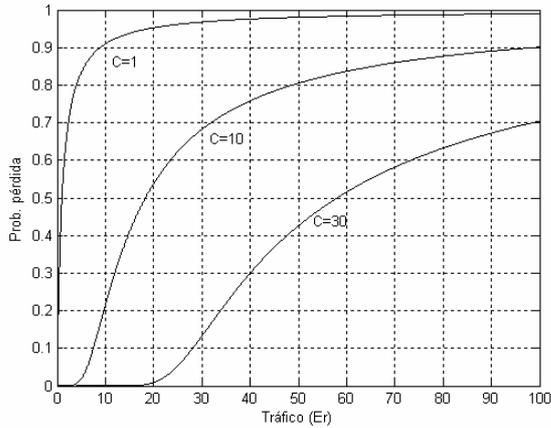
$$N = \{n_i\} \quad 0 \leq i \leq N - 1 \quad (1.4)$$

unidos por los enlaces directos

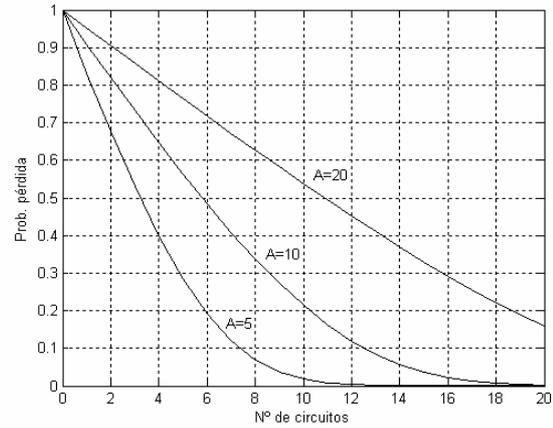
$$L = \{l_{ij}\} \quad 0 \leq i, j \leq N-1, i \neq j \quad (1.5)$$

de capacidades

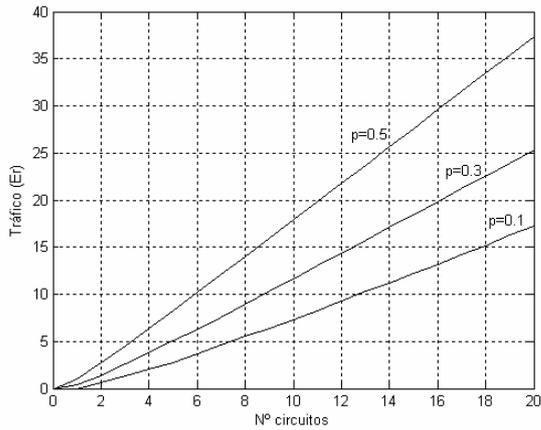
$$B = \{b_{ij}\} \quad 0 \leq i, j \leq N-1, i \neq j \quad (1.6)$$



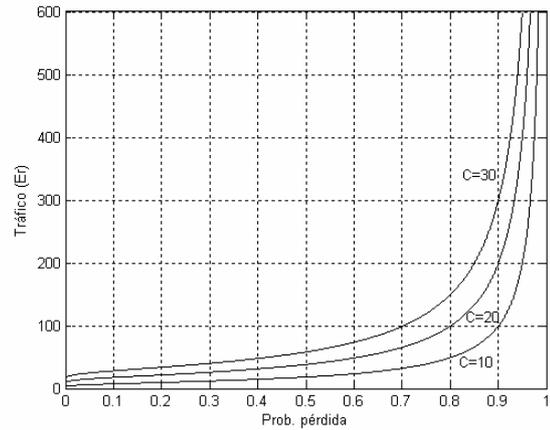
(a)



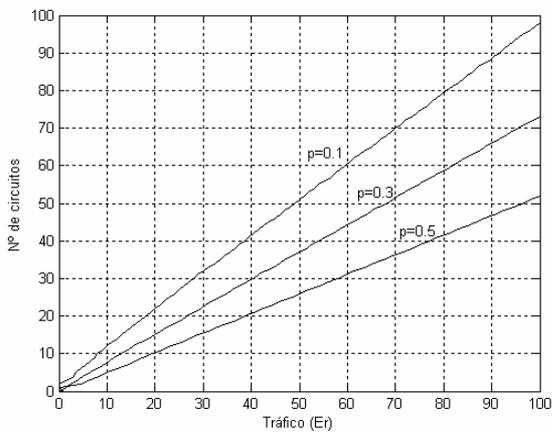
(b)



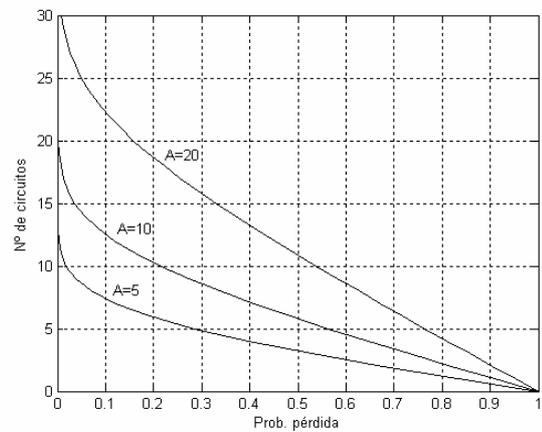
(c)



(d)



(e)



(f)

Figura 1.7 – Comportamiento Erlang-B.

En el caso de una red asimétrica, estos valores serían diferentes. La no existencia de enlace directo entre algún par de nodos es un caso particular en el que se considera la existencia de un enlace de capacidad nula.

De cara a definir el encaminamiento, se definen las rutas entre todo par de nodos como la ruta directa más las alternativas vía un nodo *tándem*; es decir,

$$R = \{R_{ij}^k\} \quad 0 \leq i, j, k \leq N - 1, i \neq j \quad (1.7)$$

$$R_{ij}^k = l_{ik} \cup l_{kj} \quad (1.8)$$

siendo k el nodo *tándem* y l_{ij} el enlace entre los nodos i y j ; se permite que k sea igual al nodo origen o el destino, definiendo así la ruta directa.

1.6. Señalización de control

En las redes de conmutación de circuitos las señales de control son el medio para gestionar la red y para establecer, mantener y finalizar las llamadas, intercambiando información entre el abonado y los conmutadores, entre los conmutadores entre sí, y entre los conmutadores y el centro de gestión de red.

1.6.1. Funciones de señalización

Las funciones de señalización más importantes son:

- Comunicación audible con el abonado, que incluye el tono de marcar, el tono de llamada, la señal de ocupado, etc.
- Transmisión del número marcado a las centrales de conmutación que intentarán establecer la conexión.
- Transmisión de información entre conmutadores, indicando que una llamada dada no se puede establecer.
- Transmisión de información entre conmutadores, indicando que una llamada ha finalizado y que la ruta puede desconectarse.
- Generación de la señal que hace que el teléfono suene.
- Transmisión de información con fines de tarifación.
- Transmisión de información indicando el estado de los equipos, las líneas para emplear en el encaminamiento, mantenimiento y diagnóstico de fallos.
- Control de equipos especiales para canales vía satélite.

A continuación se muestra un ejemplo de las señales de control entre teléfonos conectados a una misma central local:

1. Ambos teléfonos deben de estar libres (colgados) antes de la llamada. Un abonado descuelga, lo que se detecta en el conmutador de la central local a la que está conectado.
2. El conmutador responde con un tono audible de marcar, señalizando al abonado que puede marcar el número deseado.
3. El abonado que llama marca el número, lo cual se comunica al conmutador como la dirección del abonado de destino.
4. Si el abonado destinatario no está ocupado, el conmutador lo alerta acerca de la llamada entrante, enviando una señal de llamada que provoca que el teléfono suene.
5. El conmutador proporciona realimentación al abonado que llama:
 - a) Si el abonado destino no está ocupado, el conmutador devuelve un tono audible de llamada al abonado origen, mientras que simultáneamente se envía la señal de llamada al abonado llamado.

- b) Si el destino está ocupado, el conmutador envía una señal audible de ocupado al que llama.
- c) Si la llamada no puede establecerse a través del conmutador, éste envía un mensaje audible de reintento al abonado que llama.
6. El destino acepta la llamada descolgando el auricular, lo que se comunica automáticamente al conmutador.
7. El conmutador corta la señal y el tono de llamada, estableciendo una conexión entre los dos abonados.
8. La conexión se libera cuando una de las dos partes cuelga.

Cuando el abonado llamado está conectado a un conmutador diferente al del abonado origen, y ambos conmutadores están directamente conectados, son necesarias otras funciones de señalización en los enlaces que unen los dos conmutadores:

1. El conmutador origen ocupa un enlace libre entre ambos conmutadores, envía una indicación de descolgar a través del enlace y solicita un registro al otro conmutador para comunicar la dirección destino.
2. El conmutador final envía una señal de descolgar seguida por una de colgar, conocida como parpadeo. Esto indica que el registro está preparado.
3. El conmutador origen envía los dígitos de la dirección al conmutador final.

Las señales que intervienen durante la fase de señalización:

1. *Señales de supervisión* para referirnos a señales de control de carácter binario (activado/desactivado, verdadero/falso), tales como solicitud de servicio, respuesta, aviso y retorno a desocupado. Estas señales se encargan de informar acerca de la disponibilidad del abonado llamado y de los recursos de la red necesarios.
2. *Las señales de direccionamiento* que identifican al abonado. Inicialmente se genera una señal de dirección por parte de un abonado origen cuando marca un número de teléfono. La dirección resultante se puede propagar a través de la red para permitir el encaminamiento y así localizar y hacer que suene el teléfono del abonado destino.
3. *Las señales de información sobre la llamada* proporcionan al abonado información acerca del estado de la llamada. Son señales audibles y se emplean para el establecimiento y cierre de la llamada.
4. *Las señales de gestión de red* se utilizan para el mantenimiento y funcionamiento general de la red. Estas señales pueden tener forma de mensajes, como por ejemplo una lista de rutas predefinidas enviadas a una estación para la actualización de sus tablas de encaminamiento.

1.6.2. Señalización por canal común

La señalización de control tradicional en redes de conmutación de circuitos se ha realizado a través de la propia línea principal o *intracanal*. En la técnica de señalización *intracanal* se usa el mismo canal para las señales de control y la llamada propiamente dicha. Esta señalización comienza en el abonado origen y sigue la misma ruta que la llamada. Existen dos formas de señalización *intracanal*:

- *Intrabanda*, en donde se usa la misma banda de frecuencias que las señales de voz que se transmiten, con la ventaja de que al tener las señales de control las mismas propiedades electromagnéticas que las señales de voz, pueden llegar a los mismos lugares que éstas.
- *Fuera de banda*, en donde se aprovecha el hecho de que las señales de voz no utilizan todo el ancho de banda de 4 khz asignado, y se hace uso de una banda de señalización estrecha e independiente para el envío de señales de control, con la ventaja de que estas señales se pueden enviar también cuando no hay voz en el canal, lo que permite la supervisión y control continuos de la llamada.

La señalización *intra banda* tiene una serie de desventajas para las actuales redes de telecomunicaciones públicas, como son:

- La velocidad de transferencia de información se encuentra bastante limitada.
- Un canal de voz en uso sólo puede ser utilizado por las señales de control cuando no hay señales de voz en el circuito.

La señalización *fuera de banda* tiene como desventaja el pequeño ancho de banda disponible para señalización.

Otra desventaja de la señalización *intracanal* es el retardo existente desde que un abonado introduce una dirección (marca el número) hasta que la conexión se establece. Hay que reducir este retardo para las nuevas aplicaciones controladas por computador, en donde las transacciones son mensajes cortos y el tiempo de establecimiento de llamada representa una parte importante del tiempo de la transacción total.

Estas desventajas se pueden evitar mediante la *señalización por canal común*, en donde las señales de control se transmiten por rutas completamente independientes de los canales de voz, pudiendo agruparse las señales de control de varios canales de voz en un canal común. Las señales de control son mensajes que se transfieren entre los conmutadores, y entre el conmutador y el centro de gestión de red. De este modo la señalización de control de red es una red distribuida de computadores que se especializan en el transporte de mensajes de control cortos.

Hay dos modos de funcionamiento en la señalización por canal común:

- El *modo asociado*, en donde el canal común va en paralelo (asociado) a lo largo de toda la línea a los grupos de enlace entre conmutadores. Las señales de control van en canales diferentes a las señales de abonado, y dentro de un mismo conmutador las señales de control se encaminan hacia un procesador de señales de control.
- *Modo no asociado*, que es más potente pero más complejo. En este modo se hace crecer la red añadiendo puntos de conmutación especializados llamados puntos de transferencia de la señal. En este caso no existe una asignación o correspondencia, ni definitiva, ni sencilla, entre los canales de control y los grupos de enlace y como consecuencia existen dos redes independientes: la de voz y la de señalización. Con esta configuración de red se pueden establecer uno o más puntos centrales de control y puede haber un punto que haga de nodo central con una visión global del estado de la red. Este modo se usa en RDSI.

La señalización por canal común es más versátil que la señalización *intra banda* ya que es menos susceptible a posibles interferencias, obteniendo como ventajas, entre otras, una reducción del retardo en el establecimiento de la llamada.

A pesar de la mayor complejidad de la técnica de señalización por canal común, la reducción de costes en el hardware digital y el creciente carácter digital de las redes de telecomunicaciones hacen que esta técnica sea la adecuada. No obstante, siempre será necesaria la señalización *intracanal* para la comunicación con el abonado (tono de marcar, señal de indicación de llamada y señal de ocupación).

El esquema de señalización por canal común más usado es el sistema de señalización número 7 (SS7, *Signaling System number 7*), que proporciona el control interno y la inteligencia esenciales a una RDSI.

1.6.3. El protocolo SS7

En SS7 los mensajes de control se encaminan a través de la red para llevar a cabo la gestión de las llamadas (establecimiento, mantenimiento y terminación) y las funciones relativas a la gestión de red.

Estos mensajes son bloques o paquetes pequeños que se pueden encaminar a través de la red de modo que aunque la red que está siendo controlada es una red de conmutación de circuitos, la señalización de control se basa en la tecnología de conmutación de paquetes.

Los elementos de la red de señalización son:

1. Punto de señalización (SP) es un nodo de la red de señalización con capacidad de gestión de mensajes de control SS7. Un ejemplo puede ser un receptor de mensajes de control incapaz de procesar mensajes que no vayan destinados directamente a él, o un centro de control de red.
2. Un punto de transferencia de la señal (STP) es un punto de señalización capaz de encaminar mensajes de control; es decir, un mensaje recibido sobre un enlace de señalización se transfiere a otro enlace. Un ejemplo podría ser un nodo de encaminamiento puro, pudiendo realizar las funciones propias de un punto final (origen/destino) de comunicaciones.
3. Un enlace de señalización es un enlace de datos que conecta entre sí puntos de señalización.

El plano de control es el responsable del establecimiento y de la gestión de las conexiones, las cuales se solicitan por el usuario. El diálogo entre el usuario y la red se realiza entre el usuario y el conmutador local. Con este fin, el conmutador local funciona como un punto de señalización, ya que debe de llevar a cabo la conversión entre el diálogo con el usuario y los mensajes de control internos a la red, que son los que realmente realizan las acciones solicitadas por el usuario. El protocolo SS7 se usa internamente a la red para establecer y mantener una conexión dada; este proceso debe de involucrar uno o más puntos de señalización y de transferencia de señal.

Una vez que se ha establecido la conexión, la información se transfiere desde un usuario hasta el otro, extremo a extremo, en el plano de información. Para ello se establece un circuito desde el conmutador local de un usuario hasta el otro, habiéndose realizado quizás el encaminamiento a través de uno o más nodos de conmutación de circuitos, denominados centros de tránsito. Todos estos nodos (conmutadores locales, centros de tránsito) son también puntos de señalización, ya que son capaces de enviar y recibir mensajes SS7 para establecer y gestionar la conexión.

1.7. Objetivos

El Proyecto Final de Carrera que se presenta tiene como objetivo evaluar, de forma teórica y práctica, las prestaciones de los métodos de *encaminamiento dinámico no jerárquico* en redes de conmutación de circuitos. La evaluación de dichos métodos abarcará el estudio, desarrollo y simulación de los mismos. El objetivo de estos métodos es reducir al máximo la probabilidad de pérdida que percibe el usuario del servicio y, para el correcto funcionamiento de estas estrategias, en muchos casos se necesita una red de señalización asociada.

En el capítulo 2, se plantea un modelo teórico del encaminamiento dinámico en redes no jerárquicas para obtener valores teóricos de las prestaciones de la red. En este estudio teórico se demuestra que bajo ciertas condiciones aparecen márgenes de *inestabilidad*, es decir, se demuestra la existencia de más de una solución en la probabilidad de pérdida para un rango de tráfico ofrecido. Esta situación se presenta en toda red con encaminamiento dinámico no jerárquico y viene provocada por el uso abusivo de las rutas alternativas. Se demuestra de forma práctica la existencia de los distintos estados entre los que la red puede oscilar, con los resultados extraídos de las simulaciones realizadas con una herramienta programada en lenguaje C.

También, se realiza un análisis teórico del encaminamiento cuando se admite la *reserva de troncales* para uso exclusivo del tráfico directo entre cada par de nodos. En este estudio teórico se demuestra que la reserva de un número pequeño de troncales ya soluciona el problema de la *inestabilidad*. Para finalizar este capítulo, se presenta una breve reseña histórica sobre la evolución que han presentado los métodos de encaminamiento dinámico, así como una clasificación de estos métodos que aparece en recomendaciones ITU-T.

En el capítulo 3, se presenta el algoritmo de encaminamiento DAR de forma descriptiva. Con el fin de analizar de forma práctica las prestaciones de este algoritmo se presenta una herramienta de simulación programada en lenguaje C que implementa este algoritmo. A partir de los resultados obtenidos de las distintas simulaciones, se realiza una comparación de prestaciones con el encaminamiento directo, se observan las prestaciones en redes multiclase y se comprueba la influencia del tamaño de la red. Hay que indicar que las prestaciones de la red se calculan como la probabilidad de pérdida en llamadas, es decir, la probabilidad de que una llamada encuentre la ruta directa y todas las secundarias bloqueadas. Este parámetro también recibe el nombre de *grado de servicio* (GoS).

En el capítulo 4, se realiza un estudio similar al realizado en el capítulo 3, pero tomando al algoritmo de encaminamiento LLR como protagonista.

En el capítulo 5, se presenta una comparación teórica entre los distintos tipos de métodos de encaminamiento dinámico, analizando las bondades y desventajas de cada uno de ellos, y presentado las características deseadas para un algoritmo de encaminamiento dinámico. De la misma forma, a partir de los resultados de simulación obtenidos en los capítulos 3 y 4 para los métodos DAR y LLR, respectivamente, se realiza una comparación entre estos dos métodos, presentando a modo de conclusión qué situaciones son idóneas para la utilización de uno u otro método.

En el capítulo 6, se realiza un estudio sobre la aplicación que los métodos estudiados en este Proyecto Final de Carrera pueden tener en redes de conmutación de paquetes. En primer lugar, se presentan diversos casos de redes reales de conmutación de paquetes en los que se ha realizado una adaptación a esta tecnología de diversos conceptos utilizados en redes de conmutación de circuitos. Hay que decir que en las situaciones presentadas no se utilizan estos métodos para conseguir los objetivos que perseguían en conmutación de circuitos, sino que se utilizan de forma conjunta con métodos de reserva de ancho de banda en redes IP, aplicados a la gestión de tablas de encaminamiento, asociados a la selección del camino en MPLS, o aplicados al encaminamiento alternativo en redes ATM. En la segunda parte, se presenta una herramienta de simulación programada en lenguaje C, en la que se aplican los métodos de encaminamiento dinámico a un entorno de VoIP, con el mismo objetivo con el que se utilizaban en redes de conmutación de circuitos, es decir, la reducción de la probabilidad de pérdida en llamadas. En este caso, se demuestra que en un entorno de conmutación de paquetes, la utilización de estos métodos con el objetivo anteriormente mencionado es totalmente innecesaria, pues las prestaciones obtenidas con el encaminamiento directo ya superan enormemente las obtenidas en conmutación de circuitos.

En el capítulo 7, se exponen las conclusiones que se derivan del trabajo realizado.

Capítulo 2

Modelado del encaminamiento dinámico. Inestabilidades

2.1. Introducción

En este capítulo se plantea un modelo teórico del encaminamiento alternativo dinámico –o encaminamiento dinámico– en redes no jerárquicas, de forma que se puedan obtener valores teóricos de las *prestaciones de la red*, así como de otros parámetros intermedios.

La existencia de rutas alternativas a la directa provoca que haya llamadas que ocupen dos enlaces. Para valores *bajos* de tráfico, el uso de la ruta alternativa no es muy frecuente, ya que tampoco lo es el bloqueo de los enlaces directos. Cuando este poco probable bloqueo sucede, es aún mucho menos probable que coincida con un bloqueo en alguno de los dos enlaces de la(s) ruta(s) alternativa(s), por lo que la llamada en cuestión podrá ser cursada por la ruta secundaria. Este efecto se manifiesta en la red global como una disminución de la probabilidad de pérdida frente al encaminamiento directo.

Sin embargo, en condiciones de *alta* carga el bloqueo de un enlace es muy probable, por lo que serán muchas las llamadas que intenten las rutas alternativas. Cuando aumenta el número de llamadas en curso que emplean rutas secundarias, el efecto que se produce es el de un aumento de la probabilidad de pérdida frente al encaminamiento directo. Además, en esta situación, el tráfico real cursado disminuye, ya que las llamadas que emplean dos enlaces *generan* el doble de tráfico, produciéndose como consecuencia un aumento en la probabilidad de pérdida de los enlaces. Por otra parte, también en condiciones de alta carga puede suceder que la mayor parte de las llamadas se cursen por los enlaces directos, debido a la aleatoriedad del tráfico ofrecido, aumentando así el número de llamadas cursadas y disminuyendo la probabilidad de pérdida. Para valores altos de tráfico aparecen, pues, varios estados posibles, en principio, entre los que la red puede *oscilar*. La existencia de más de una solución en la probabilidad de pérdida para un rango de tráfico ofrecido recibe el nombre de *inestabilidad*. Esta *inestabilidad*, que presenta toda red con encaminamiento dinámico no jerárquico, no supondría ningún problema si dichas soluciones permaneciesen dentro de valores tolerables de probabilidad de pérdida. Sin embargo, en el estado de alta probabilidad de pérdidas, ésta es mucho mayor que la que se obtiene con encaminamiento directo, por lo que el GoS (*Grado de Servicio*) medio empeora.

Tras el desarrollo expuesto parece, en principio, que el origen de la *inestabilidad* puede estar en el sobreuso que se puede hacer de las rutas alternativas. La solución a esta situación pasa por la utilización de *reserva de troncales*. A continuación se demostrará la existencia de las *inestabilidades* de forma teórica y se comprobará posteriormente mediante simulación.

2.2. Modelo EFP

2.2.1. Modelado general

Para hacer un modelado del encaminamiento alternativo dinámico, en principio independiente del algoritmo concreto, se empleará el modelo EFP (*Erlang Fixed Point*):

- Se supone que el tráfico es de Poisson, tanto el directo como el desbordado.
- Las probabilidades de bloqueo de los enlaces se consideran independientes.

La notación empleada usa las letras a , c y p para el tráfico, número de troncales entre nodos y probabilidad de bloqueo, respectivamente. En el caso del tráfico y de la probabilidad de bloqueo aparecen letras minúsculas (a , p) cuando se hace referencia a tráfico y probabilidad de bloqueo *interiores*, en un enlace determinado, mientras que se usan las mayúsculas (A , P) para el tráfico y la probabilidad de bloqueo *exteriores*, entre pares de nodos. Los subíndices hacen referencia a los nodos que intervienen.

La suposición de que el tráfico es de Poisson permite emplear la fórmula Erlang-B para calcular la probabilidad de bloqueo del enlace que une los nodos i y j como

$$p_{ij} = B(a_o^{ij}, c_{ij}) \quad (2.1)$$

en función del tráfico ofrecido al enlace a_o^{ij} y el número de circuitos del mismo c_{ij} . Por lo tanto, el tráfico cursado por el enlace es

$$a_c^{ij} = a_o^{ij} (1 - p_{ij}) \quad (2.2)$$

Las *prestaciones de la red* se van a calcular como la probabilidad de pérdida en llamadas, es decir, la probabilidad de que una llamada encuentre la ruta directa y todas las secundarias bloqueadas. Este parámetro también recibe el nombre de *grado de servicio* (GoS). La hipótesis de independencia entre probabilidades de bloqueo, permite calcular la probabilidad de pérdida en llamadas como el producto del bloqueo de todas las rutas, tanto directas como alternativas:

$$P_{ij} = p_{ij} \prod_{k \in M_{ij}} [1 - (1 - p_{ik})(1 - p_{kj})] \quad (2.3)$$

siendo M_{ij} el conjunto de nodos *tándem* para el par (i, j) . Por lo tanto, el tráfico cursado entre cada par de nodos es

$$A_c^{ij} = A_o^{ij} (1 - P_{ij}) \quad (2.4)$$

El tráfico cursado por cada enlace, además de la expresión (2.2), puede expresarse como suma de tres términos: el tráfico directo cursado, más los tráficos cursados en los dos sentidos cuando el enlace en cuestión es uno de los dos de una ruta alternativa, teniendo en cuenta que los enlaces son bidireccionales,

$$a_c^{ij} = A_o^{ij} (1 - p_{ij}) + \sum_{\forall k / i \in M_{kj}} a_c^{kij} + \sum_{\forall k / j \in M_{ki}} a_c^{kji} \quad (2.5)$$

siendo a_c^{kij} el tráfico cursado entre los nodos k y j vía el nodo i , que puede expresarse en función del tráfico *exterior* ofrecido al par de nodos y la probabilidad de pérdida entre ellos, de la forma

$$a_c^{kij} = A_o^{kj} p_{kj} \left\{ 1 - \prod_{l \in M_{kj}} [1 - (1 - p_{kl})(1 - p_{lj})] \right\} \pi_{kij} \quad (2.6)$$

con π_{kij} la probabilidad de uso del nodo i como nodo *tándem* del par (k, j) .

2.2.2. Particularización a una red simétrica totalmente conectada

La resolución de las ecuaciones anteriores es suficientemente compleja como para considerar diferentes topologías de red. Se considerarán redes simétricas totalmente conectadas para diferentes valores de número de nodos, circuitos por enlace o tráfico entre cada par de nodos. Para estas redes, las ecuaciones (2.1) a (2.6) pueden simplificarse. Eliminando los subíndices que hacen referencia a los nodos involucrados, y operando, se obtiene

$$p = B(a_o, c) \quad (2.7)$$

$$a_c = a_o(1 - p) \quad (2.8)$$

$$P = p[1 - (1 - p)^2]^M \quad (2.9)$$

$$A_c = A_o(1 - P) \quad (2.10)$$

$$a_c = A_o(1 - p) + 2Ma_c^{(a)} \quad (2.11)$$

$$a_c^{(a)} = A_o p \left(1 - \frac{P}{p}\right) \frac{1}{M} \quad (2.12)$$

siendo $a_c^{(a)}$ el tráfico cursado por un enlace cuando éste actúa como parte de una ruta alternativa, y M el número de rutas secundarias para cada par de nodos, es decir $M=N-2$. Sustituyendo (2.12) en (2.11)

$$a_c = A_o(1 + p - 2P) \quad (2.13)$$

De (2.8) y (2.13)

$$a_o = A_o \frac{1 + p - 2P}{1 - p} \quad (2.14)$$

Agrupando las ecuaciones (2.7), (2.9) y (2.14)

$$p = B(a_o, c) \quad (2.15a)$$

$$P = p[1 - (1 - p)^2]^M \quad (2.15b)$$

$$a_o = A_o \frac{1 + p - 2P}{1 - p} \quad (2.15c)$$

Sustituyendo (2.15b) en (2.15c) y el resultado en (2.15a)

$$p = B\left(\frac{A_o}{1 - p} \left\{1 + p - 2p[1 - (1 - p)^2]^M\right\}, c\right) \quad (2.16)$$

que es una ecuación no lineal en p ya que B es la función Erlang-B. No puede, pues, encontrarse una solución analítica para p , habiendo de recurrir a métodos numéricos. Nótese que los datos son el tráfico ofrecido A_o , el número de rutas alternativas M y el número de troncales entre todo par de nodos c .

2.3. Resolución iterativa de la ecuación de red

La ecuación (2.16) se escribirá de forma resumida como

$$p = B[a_o(p), c] \tag{2.17}$$

Resolverla equivale a encontrar los ceros de la función

$$f(p) = p - B[a_o(p), c] \tag{2.18}$$

En primer lugar, se demostrará que $f(p)$ tiene al menos una raíz en $[0,1]$. Evaluando en $p=0$ y $p=1$ y teniendo en cuenta que la función Erlang-B devuelve un valor en ese mismo intervalo, que el tráfico ofrecido es finito y que el número de troncales es no nulo, se tiene que

$$f(0) = 0 - B[a_o(0), c] = -B[A_o, c] < 0 \tag{2.19a}$$

$$f(1) = 1 - B[a_o(1), c] = 1 - B[A_o, c] > 0 \tag{2.19b}$$

tras calcular el límite cuando p tiende a 1 por L'Hôpital. En virtud del teorema de Bolzano, a partir de 2.19a y 2.19b, y dado que $f(p)$ es una función continua en su dominio, se deduce que dicha función tiene un número impar de soluciones (y por lo tanto, existe al menos una solución).

En la figura 2.1 se ha representado $f(p)$ de forma gráfica para tres valores diferentes de tráfico ofrecido entre cada par de nodos de la red y en la tabla 2.1 aparecen sus raíces. Aunque se estudiará a continuación con más detalle, se observa que hay tres intervalos de tráfico en los que el comportamiento es distinto. Para tráfico bajos (<80 Er, nótese que $c=100$), la probabilidad de bloqueo de los enlaces es pequeña. Para tráfico un poco mayores (85 Er), aparecen tres posibles soluciones para la ecuación planteada. Para tráfico aún mayores (>90 Er), vuelve a obtenerse sólo una solución, la de alta probabilidad de bloqueo. En el intervalo en el que hay tres soluciones, la red va a oscilar entre los tres estados. Quedaría así demostrado el comportamiento *inestable* de la red en determinadas circunstancias.

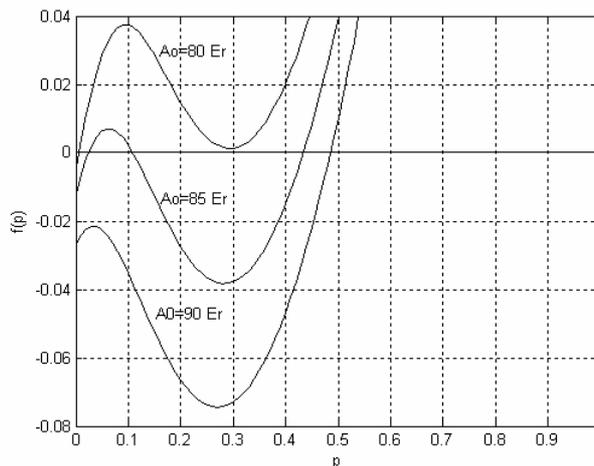


Figura 2.1 – Aspecto de $f(p)$ para diferentes valores de tráfico ofrecido. Red simétrica, totalmente conectada, $N=5$, $M=3$, $c=100$.

$A_o(Er/c)$	Solución 1	Solución 2	Solución 3
0,8	0,005		
0,85	0,025	0,110	0,435
0,9			0,485

Tabla 2.1 – Raíces de la figura 2.1.

Aunque se dispone de una expresión analítica para la ecuación que se desea resolver, no se puede despejar la variable p . En este proyecto se ha optado por la realización de un programa MATLAB, que obtenga los valores de la función $f(p)$ para sucesivos valores del tráfico ofrecido entre cada par de nodos de la red (A_o). A continuación obteniendo los ceros de la función $f(p)$, se consiguen los valores de la probabilidad de bloqueo que le corresponden al tráfico ofrecido en cuestión. De esta forma se obtiene la probabilidad de bloqueo que presenta un enlace de la red (p), en función del tráfico ofrecido al mismo.

Ya que, como puede verse en la ecuación (2.18), la ecuación $f(p)$ requiere la resolución de la función de Erlang-B, para simplificar dicha resolución y disminuir la carga computacional se ha optado por la utilización de la expresión recursiva para la función Erlang-B, que tiene la forma

$$\frac{1}{B(a_o, c+1)} = 1 + \frac{1}{B(a_o, c)} \cdot \frac{c+1}{a_o} \quad (2.20)$$

y que toma como valor inicial

$$B(a_o, 1) = \frac{a_o}{1 + a_o} \quad (2.21)$$

En la figura 2.2 se presenta el mencionado programa MATLAB, utilizado para la resolución de la ecuación de red y la obtención de la probabilidad de bloqueo de un enlace en función del tráfico ofrecido al mismo.

```

k=1; %Índice para cada valor de tráfico ofrecido.
for A=0:0.01:10; %Se calcula el valor de la función f(p) para valores del tráfico ofrecido entre 0 y 10.
    %Cada valor de A va a representar una función f(p) distinta.
    h=1; %Índice para cada valor la probabilidad p.
    for p=0:0.001:1 %Valores posibles de la probabilidad p.
        a0=(A/(1-p))*(1+p-2*p*(1-(1-p)^2)^M); %Tráfico ofrecido en función del valor de p.
        Binv=(a0+1)/a0; %Representa la inversa del valor de B(a0,1).
        for j=2:1:c %Ecuación recursiva para encontrar B(a0,c).
            Binv=1+Binv*j/a0;
        end
        f(h)=p-1/Binv; %Representa f(p).
        h = h+1;
    end

    x=0:0.001:1; %Valores posibles de la probabilidad p.
    indices=find(f<0.001 & f>-0.001) %Buscamos el valor de los ceros de la función f(p).
    P(k)=x(indices(1)) %Los ceros de f(p) representan la probabilidad de bloqueo,
    %para el tráfico ofrecido correspondiente.

    k=k+1;
end;

```

Figura 2.2 – Programa MATLAB para la resolución de la ecuación de red.

Como se puede observar en la figura 2.2, el valor obtenido para la probabilidad de bloqueo representa una aproximación al valor real, ya que resulta computacionalmente intratable la obtención del valor exacto de dicha probabilidad de bloqueo. Sin embargo, se puede decir que el valor obtenido converge hacia la solución real con una precisión suficientemente elevada de acuerdo al objetivo que perseguimos (la representación gráfica de los valores de la probabilidad de bloqueo, para realizar un estudio comparativo en función de los distintos parámetros que intervienen en la ecuación de red).

2.4. Resultados teóricos

Se ofrecen los resultados gráficos obtenidos a partir del desarrollo expuesto. Son muy interesantes a la hora de analizar el comportamiento interno de la red cuando se permiten rutas alternativas.

2.4.1. Red estable

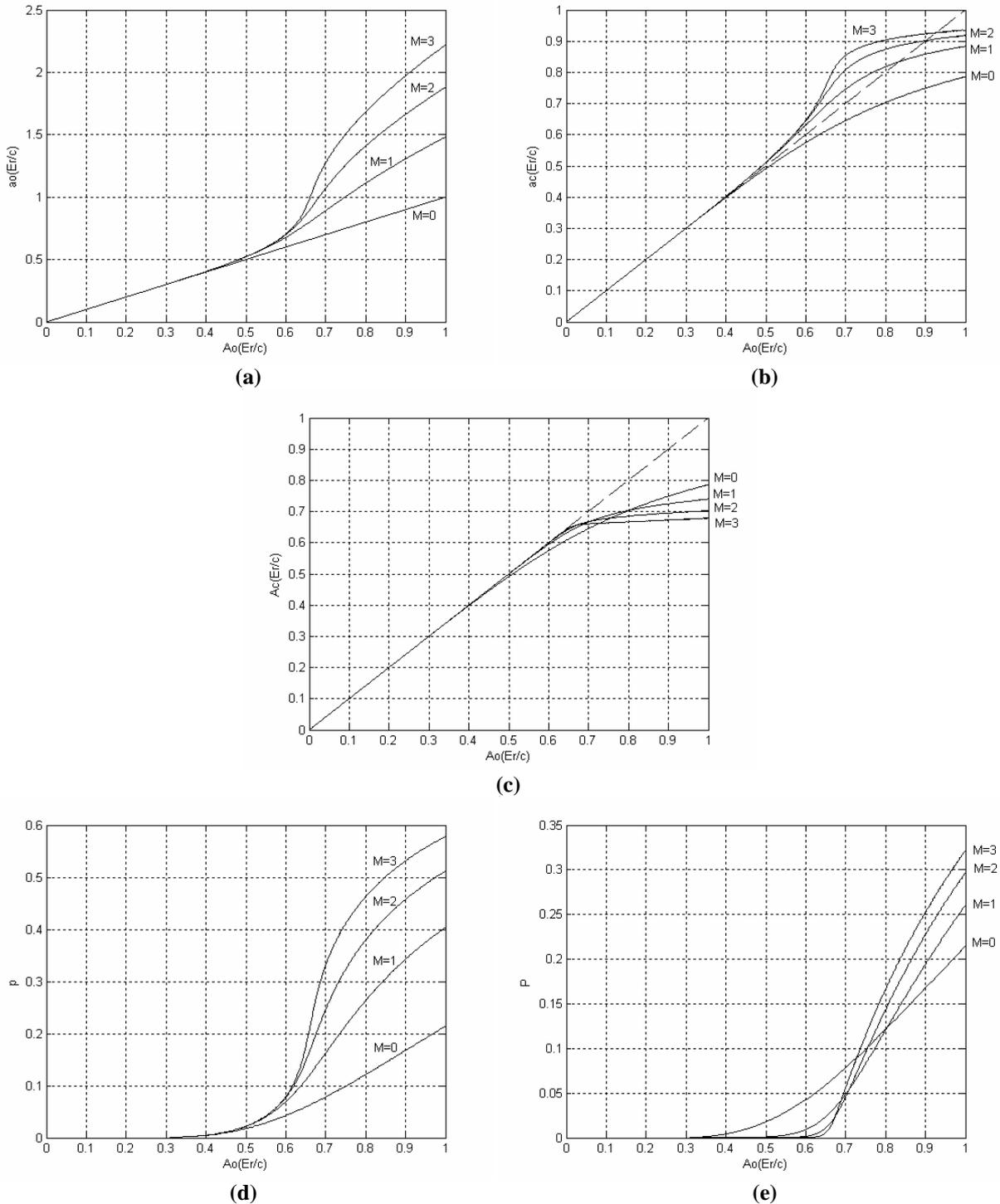


Figura 2.3 – Comportamiento de una red simétrica totalmente conectada ($N=5$, $c=10$) frente al tráfico normalizado ofrecido entre cada par de nodos para diferente número de rutas alternativas (M). **(a)** Tráfico normalizado ofrecido a cada enlace. **(b)** Tráfico normalizado cursado por cada enlace. **(c)** Tráfico normalizado cursado por la red entre cada par de nodos. **(d)** Probabilidad de bloqueo de un enlace. **(e)** Probabilidad de bloqueo de ruta.

Los resultados de las figuras 2.3a a 2.3e se refieren a una red simétrica totalmente conectada de 5 nodos y de 10 circuitos por enlace, para diferente número de rutas alternativas. Se considera una red pequeña, pero que ya presenta claramente algunos de los efectos del encaminamiento.

En la figura 2.3a se representa el tráfico ofrecido a cada enlace de la red frente al que se ofrece a cada par de nodos. La curva es una recta de pendiente unidad para $M=0$, caso del encaminamiento directo, como era de esperar. Cuanto mayor es el número de rutas secundarias de que se dispone, mayor es el tráfico ofrecido a cada enlace ya que, además del directo, cada enlace soporta el tráfico desbordado del resto de los enlaces. Para tráficos bajos todas las curvas coinciden; es decir, todo el tráfico se cursa por el enlace directo, hecho lógico ya que la probabilidad de bloqueo será baja por lo que la ruta directa estará libre.

En la figura 2.3b se ha representado el tráfico cursado por los enlaces. La línea discontinua marca la recta de pendiente unidad, es decir, el tráfico que cursaría una red de recursos infinitos. Análogamente, para poca carga todo el tráfico ofrecido entre nodos es cursado. Para cargas un poco mayores (entorno a los 0.7 Erlangs/circuito) la red con encaminamiento alternativo cursa más tráfico que el ofrecido, debido a la ocupación de dos enlaces por parte de algunas llamadas. Para alta carga, la red entra en una zona de saturación donde prácticamente ya no aumenta el tráfico cursado debido al incremento de la probabilidad de pérdida.

El hecho anterior es mucho más claro a partir del análisis de la figura 2.3c, donde se representa el tráfico cursado entre pares de nodos. En ningún momento se cursa más tráfico del ofrecido ya que este tráfico externo contabiliza llamadas en curso y no enlaces ocupados, como en la gráfica anterior. Se pueden diferenciar claramente las tres zonas anteriores. Para tráficos bajos, todos los algoritmos cursarían todo el tráfico, por lo que el encaminamiento directo sería el deseable dada su menor complejidad. También sería la opción correcta para cargas elevadas, precisamente por la limitación en el número de enlaces que puede ocupar una llamada. Hay, sin embargo, una zona de carga intermedia donde el tráfico cursado por las redes con rutas alternativas es mucho mayor que el directo. También hay diferencias entre ellas ya que cuanto mayor es el valor de M , mayor es el tráfico para el que la curva real se separa de la ideal, pero mayor es también el efecto de saturación. Por lo tanto aparecen intervalos en los que las mejores prestaciones se obtienen para diferentes valores de M . Esto se observa mejor en la ampliación mostrada en la figura 2.4.

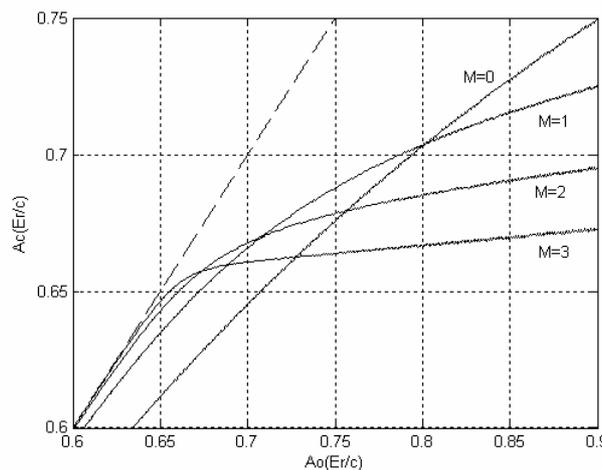


Figura 2.4 – Ampliación de la figura 2.3c.

En la figura 2.3d se representa la probabilidad de bloqueo de un enlace, probabilidad que se incrementa a medida que lo hace el tráfico ofrecido, siendo tanto mayor cuantas más rutas alternativas disponga cada par de nodos. La curva de $M=0$ coincide con la Erlang-B, analizada en el capítulo anterior.

Frente a la probabilidad de bloqueo interna anterior, la figura 2.3e muestra la probabilidad de bloqueo de ruta, es decir, la probabilidad de que tanto el enlace directo como los secundarios se encuentren bloqueados. Se trata, pues, de un concepto externo, por lo que será tomado como indicador de las *prestaciones de la red*, como ya se indicó anteriormente. Una vez más, la curva de $M=0$ es la Erlang-B. Para cada valor de M puede encontrarse un valor máximo de tráfico para el que existe ganancia por encaminamiento; más allá de ese valor, la degradación por tolerar la ocupación de varios enlaces por una llamada es mayor que el beneficio de poseer rutas alternativas; esta ganancia existe para cargas bajas y medias, cuando el bloqueo de la ruta directa es poco o medianamente probable y la ocupación de dos enlaces por una llamada no demasiado frecuente. Por otra parte, también se observa que el umbral a partir del cual aparece la probabilidad de bloqueo de ruta es mayor para mayores valores de M , debido a que se dispone de más rutas alternativas para enfrentarse a bloqueos esporádicos de otras rutas, aunque su crecimiento tras este umbral es más pronunciado, por un mayor número de rutas ocupadas, apareciendo de nuevo intervalos en los que cada valor de M es óptimo, como se observa de forma detallada en la ampliación mostrada en la figura 2.5.

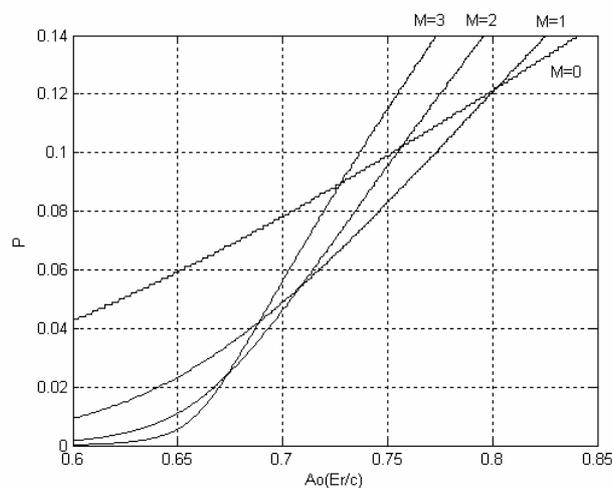


Figura 2.5 – Ampliación de la figura 2.3e.

2.4.2. Desestabilización por número de rutas alternativas

En este apartado se demostrará que hay redes *condicionalmente estables* en función del número de rutas alternativas que cada par de nodos posee. Así, se verá que a partir de un cierto valor de M aparecen márgenes de inestabilidad. La red considerada es también de 5 nodos, pero con enlaces de 100 circuitos de capacidad.

Los resultados del estudio teórico se muestran en las gráficas 2.6a a 2.6e. Se trata de las mismas gráficas del apartado anterior, por lo que poseen las mismas características anteriormente comentadas. Sin embargo, aparecen unos patrones de comportamiento nuevos que no se tenían en la red del apartado anterior. Así, se observa que para $M=2$ y $M=3$ hay márgenes de tráfico ofrecido entre nodos para los que la red puede presentar tres soluciones en el parámetro representado, es decir, aparecen márgenes de *inestabilidad*. Nótese cómo el aumento del número de rutas secundarias provoca una curvatura de la gráfica considerada, partiendo de la correspondiente al encaminamiento directo, hasta que aparecen las inestabilidades mencionadas. Por ese motivo se ha hablado de redes *condicionalmente estables*, en el sentido de que lo serán dependiendo de las rutas secundarias que el algoritmo de encaminamiento considere. Por lo tanto, puede demostrarse que, bajo determinadas condiciones, las prestaciones de una red se incrementan al restringir el número de rutas alternativas.

Como se ha comentado anteriormente, el verdadero problema de la desestabilización de las redes al permitir el *encaminamiento dinámico no jerárquico* es que alguna de las soluciones corresponda a estados de alto GoS (probabilidad de bloqueo *exterior*, P).

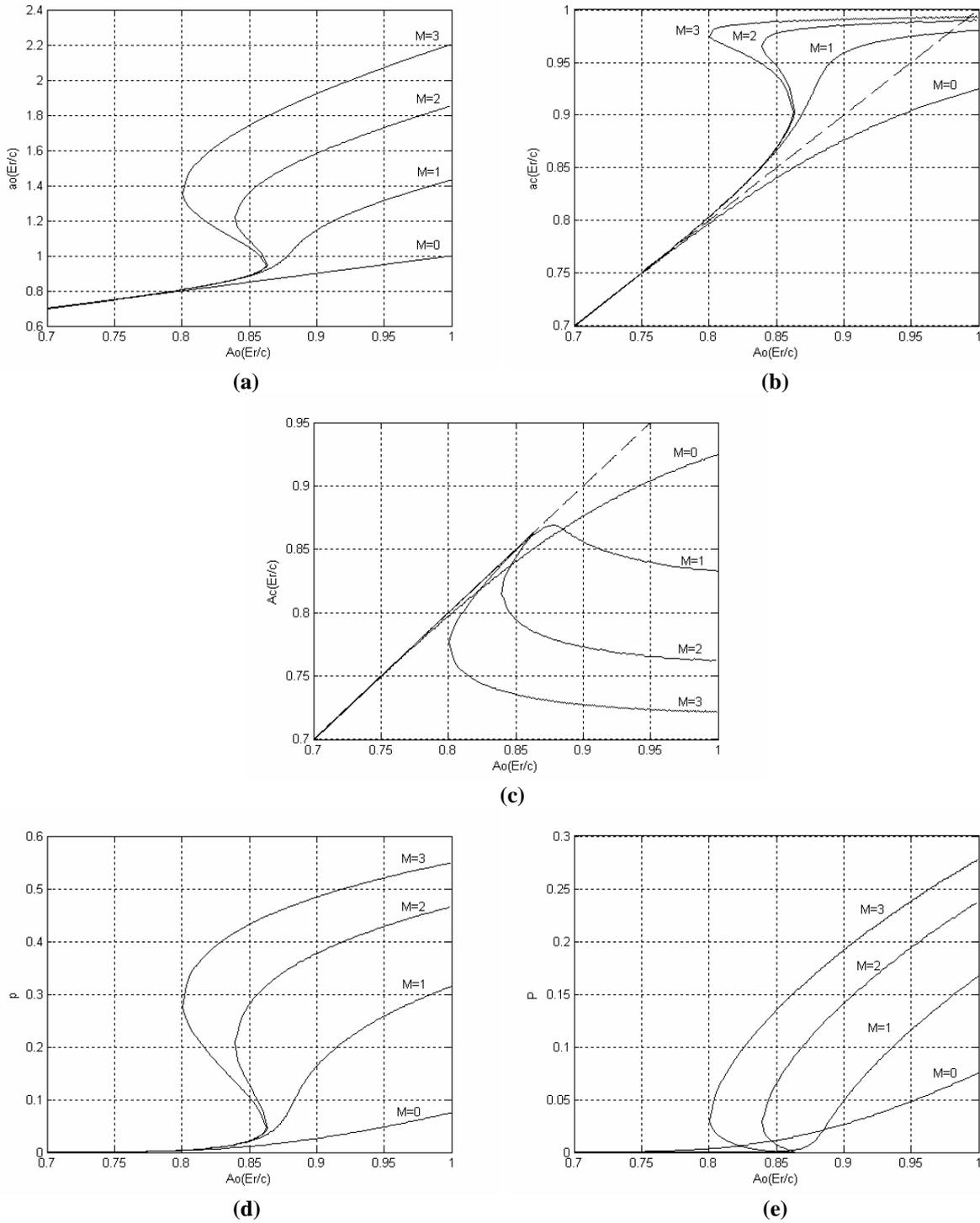


Figura 2.6 – Comportamiento de una red simétrica totalmente conectada ($N=5$, $c=100$) frente al tráfico normalizado ofrecido entre cada par de nodos para diferente número de rutas alternativas. **(a)** Tráfico normalizado ofrecido a cada enlace. **(b)** Tráfico normalizado cursado por cada enlace. **(c)** Tráfico normalizado cursado por la red entre cada par de nodos. **(d)** Probabilidad de bloqueo de un enlace. **(e)** Probabilidad de bloqueo de ruta.

En la figura 2.6e se observa cómo las soluciones de alto GoS superan con mucho los valores máximos tolerables del mismo, un 1% habitualmente, degradando así enormemente las prestaciones de la red. Así, en el ejemplo considerado, para $M=3$ y un tráfico de 85 Erlangs, los tres posibles GoS son de 0,0256%, 0,0934% y 13,62%, este último a todas luces inaceptable.

2.4.3. Desestabilización por capacidad en los enlaces

En el apartado anterior se ha comprobado que las redes que se están tratando son condicionalmente estables dependiendo del valor del parámetro M . Sin embargo, se ha podido ver que la misma red de 5 nodos es incondicionalmente estable para enlaces de 10 circuitos y no lo es para enlaces de 100 circuitos. Por lo tanto, la velocidad de los enlaces será un parámetro relevante a la hora de estudiar la inestabilidad de la red. Así, el siguiente ejemplo considera una red de 5 nodos, en la que cada par de nodos de la misma sólo dispone de una ruta alternativa, $M=1$, y se varía el número de circuitos. Los resultados se muestran en las figuras 2.7a a 2.7e, figuras correspondientes a los cinco parámetros representados en los apartados anteriores.

Se observa la condicionalidad de la estabilidad, apareciendo la inestabilidad a partir de una determinada capacidad de los enlaces, entorno a los 400 circuitos por enlace. Parece que se refuerza la hipótesis planteada acerca del sobreuso de rutas alternativas como origen de la inestabilidad, ya que ésta aparece tanto para redes o algoritmos que consideran más rutas alternativas, como para redes de relativa alta capacidad.

2.4.4. Número medio de enlaces ocupados

Al comienzo del capítulo, y de forma intuitiva, se justificó la desestabilización de las redes con encaminamiento alternativo dinámico por la posibilidad de que las llamadas cursadas ocupasen mayoritariamente bien rutas alternativas (de dos enlaces) o bien enlaces directos. Se verá a continuación que los tres estados de GoS que pueden aparecer para ciertos márgenes de tráfico ofrecido se corresponden con tres estados en el número medio de enlaces ocupados por las llamadas en curso. Para calcular este número basta comparar el tráfico *real* en curso con el *aparente*, que considera el tráfico de una llamada cursada por una ruta alternativa como doble. En definitiva, hay que comparar el tráfico cursado entre un par de nodos, A_c o tráfico *real* o *exterior*, y el tráfico cursado por un enlace, a_c o tráfico *aparente* o *interior*, que contabiliza doble el tráfico alternativo. Así, el número medio de enlaces ocupados vendrá dado por el cociente entre ambos tráficos; es decir,

$$e = \frac{a_c}{A_c} \quad (2.22)$$

El parámetro anterior está relacionado con el número medio de llamadas que se cursan por una ruta alternativa de dos enlaces que se puede calcular como

$$n_a = e - 1 = a_c / A_c - 1 \quad (2.23)$$

En la figura 2.8 se ha representado el número medio de enlaces ocupados para diferentes valores de M . Al igual que se ha comprobado para los parámetros ya estudiados, en un determinado margen de tráfico y a partir de un cierto número de rutas alternativas, aparecen varias soluciones en este parámetro. Lo más interesante es comprobar que existe una correspondencia entre las soluciones de este parámetro y las del GoS de forma proporcional, es decir, el estado de mayor número medio de enlaces ocupados se corresponde con el de mayor GoS, demostrando la relación causa-efecto que se ha propuesto como explicación. Cuando $M=0$, las llamadas en curso siempre usan un enlace, ya que el encaminamiento es directo.

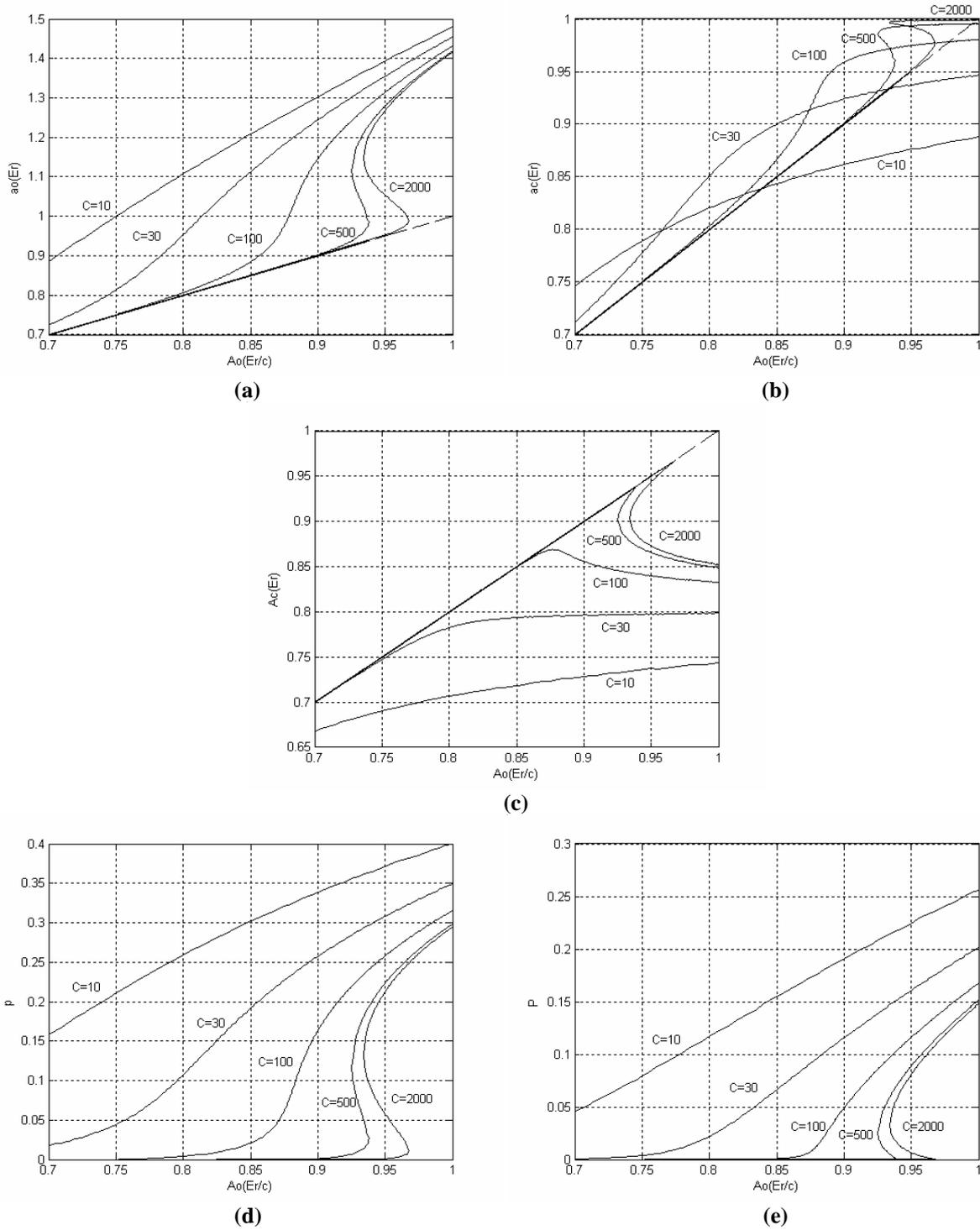


Figura 2.7 – Comportamiento de una red simétrica totalmente conectada ($M=1$) frente al tráfico normalizado ofrecido entre cada par de nodos para enlaces de diferente capacidad. **(a)** Tráfico normalizado ofrecido a cada enlace. **(b)** Tráfico normalizado cursado por cada enlace. **(c)** Tráfico normalizado cursado por la red entre cada par de nodos. **(d)** Probabilidad de bloqueo de un enlace. **(e)** Probabilidad de bloqueo de ruta.

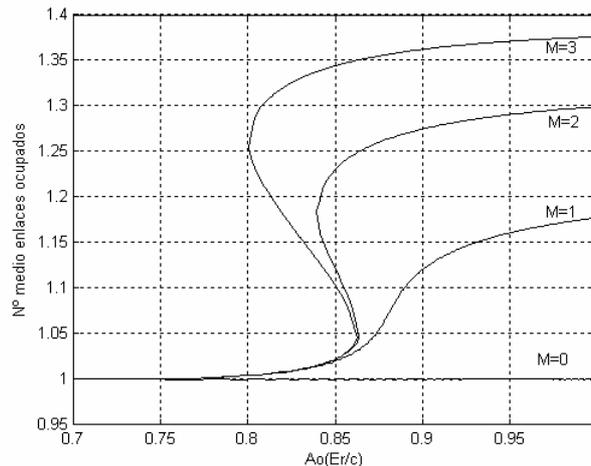


Figura 2.8 – Número medio de enlaces ocupados en una red simétrica totalmente conectada ($N=5$, $c=100$) frente al tráfico normalizado ofrecido entre cada par de nodos para diferente número de rutas alternativas.

2.5. Resultados de simulación

Si bien ha quedado probado de forma teórica que el encaminamiento dinámico en redes no jerárquicas presenta problemas de inestabilidad, es interesante comprobar que la red oscila entre los diferentes estados solución de la ecuación de red. Ello puede comprobarse sobre cualquiera de los parámetros estudiados, si bien se tomará el GoS por ser la medida elegida para las prestaciones de la red.

Para ello, se ha realizado la programación de un simulador en lenguaje C, en el que el algoritmo de búsqueda de la ruta alternativa es el algoritmo *aleatorio puro*. El primer intento de cursar la llamada siempre es a través del enlace directo entre los nodos que se desea conectar. En caso de que éste esté bloqueado, se elige de forma totalmente aleatoria una ruta alternativa entre todas las posibles. Si ésta también está bloqueada, se elige aleatoriamente otra ruta alternativa entre las restantes, y así sucesivamente. Una llamada será rechazada si todas las rutas alternativas están bloqueadas.

Por otra parte, de forma adicional, y con el fin de mejorar el procesamiento de los resultados obtenidos de las distintas simulaciones realizadas, se ha realizado un programa AWK. Este programa nos permite procesar las trazas de las distintas simulaciones, modificando el tiempo de observación en el que se contabilizan los distintos eventos de interés. En la figura 2.9 se muestra dicho programa AWK.

Antes de proceder a la observación de los problemas de inestabilidad, vamos a comprobar los resultados que se obtienen en una red que, a partir de los resultados teóricos, podemos verificar que es incondicionalmente estable. En la figura 2.10 se muestran los resultados obtenidos en una simulación de una red simétrica totalmente conectada con enlaces de 10 circuitos ($c=10$), permitiendo la utilización de 3 rutas alternativas ($M=3$), y con un tráfico ofrecido de 8 Erlangs. En esta figura se puede apreciar como el GoS oscila entorno a un valor fijo de 0,16, aproximadamente, a lo largo de los 1500 segundos que dura la simulación. Por otra parte, si observamos los resultados teóricos obtenidos en la figura 2.3e, se aprecia que para una red de las características anteriormente mencionadas, y un tráfico ofrecido de 8 Erlangs, se obtiene un GoS de 0,16, aproximadamente. Por lo tanto, resulta lógico que en la evolución temporal de este parámetro, observado en la figura 2.10, se observen oscilaciones entorno al valor de 0,16, ya que para un tiempo de simulación suficientemente elevado el valor medio de esta evolución temporal converge, teóricamente, al valor de 0,16.

Otra característica importante de la simulación de una red estable es que, a diferencia de las redes inestables, todos los enlaces de la red convergen al valor teórico correspondiente (0,16, en el caso anterior). Por su parte, en las redes inestables, tal y como se comentará posteriormente, los enlaces de una misma red no convergen al final a un mismo valor medio del GoS.

```
#script
awk -f codigol.awk -v intervalo=27 -v fichero=llegada.txt perdida.txt llegada.txt > aux.txt
awk -f codigo2.awk -v intervalo=27 aux.txt > matlab.mat

#codigol.awk
BEGIN{ global=1; local=0; i=0}
{if(FILENAME==fichero){
if(i==0){
printf "%d ",local;
if(int(2500/intervalo)==2500/intervalo){rondas=2500/intervalo;}
else {rondas=int(2500/intervalo)+1;}
while(global<rondas){
printf "0 ";
global+=1;}
global=1; local=0; i+=1;}}
while($1>global*intervalo){
printf "%d ",local;
local=0;
global+=1;}
local+=1;
}
}
END{ printf "%d ",local;
if(int(2500/intervalo)==2500/intervalo){rondas=2500/intervalo;}
else {rondas=int(2500/intervalo)+1;}
while(global<rondas){
printf "0 ";
global+=1;}
}

#codigo2.awk
BEGIN{
if(int(2500/intervalo)==2500/intervalo){rondas=2500/intervalo;}
else {rondas=int(2500/intervalo)+1;}
{
for(i=1;i<=rondas;i++){
printf "%f\n",$(i)/(i+rondas);
}
}
}
```

Figura 2.9 – Programa AWK para procesamiento de resultados.

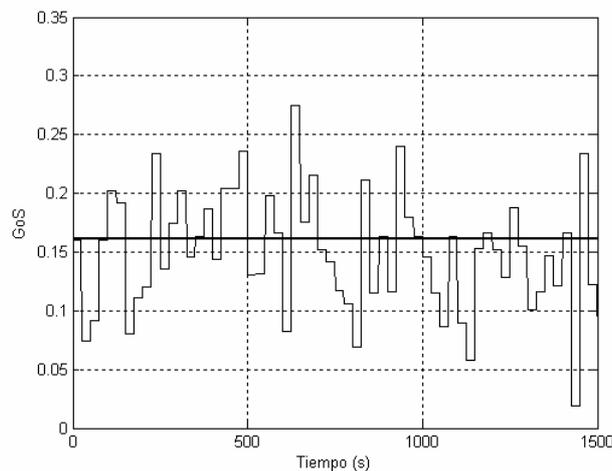


Figura 2.10 – Evolución temporal del GoS de una red simétrica totalmente conectada ($N=5$, $c=10$, $M=3$).

En la figura 2.11 se observan los resultados obtenidos en una simulación de una red simétrica totalmente conectada con enlaces de 100 circuitos ($c=100$), y permitiendo la utilización de 3 rutas alternativas ($M=3$). En esta figura se observa de forma muy clara que la red oscila entre los tres diferentes estados, con GoS aproximados de 7.5%, 1% y prácticamente nulo para el de bajo GoS. Es importante señalar que los picos en GoS, tanto los de 1% como los de 7.5%, se producen para ráfagas sostenidas en el tráfico ofrecido superiores a la media.

En los resultados teóricos obtenidos para la red anterior, mostrados en la figura 2.6e, se observa que para un tráfico ofrecido de 85 Erlangs se obtienen tres soluciones para el GoS. Estas soluciones se corresponden con los valores aproximados de 7.5%, 1% y prácticamente nulo. Por lo tanto, queda demostrado de forma práctica que en determinadas circunstancias aparecen inestabilidades, en redes que utilizan encaminamiento dinámico no jerárquico.

Hay que decir, sin embargo, que en este caso, para un tiempo de simulación suficientemente elevado, la media del GoS a lo largo del tiempo no va a coincidir con ninguna de las soluciones obtenidas de forma teórica, sino que se corresponderá con un valor “medio” entre las 3 soluciones posibles. Y, tal como se adelantó anteriormente, este valor “medio” no va a coincidir en todos los enlaces, debido a la aleatoriedad con la que se producen las inestabilidades en cada uno de ellos.

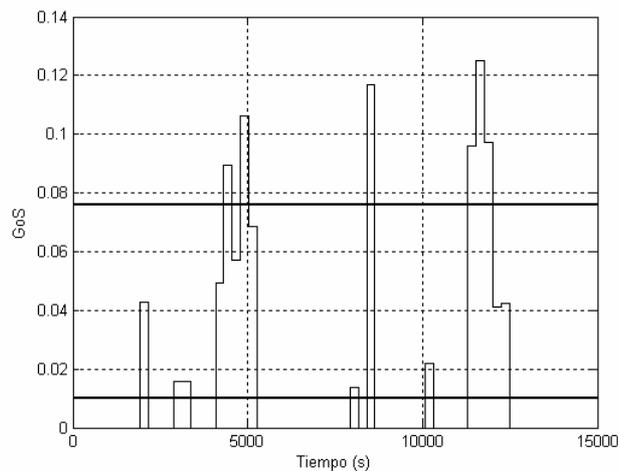


Figura 2.11 – Evolución temporal del GoS de una red simétrica totalmente conectada ($N=5, c=100, M=3$).

2.6. Método de reserva de troncales

2.6.1. Desarrollo teórico

En este apartado se realiza un análisis teórico del encaminamiento cuando se admite la reserva de troncales para uso exclusivo del tráfico directo entre cada par de nodos. Cuando se utiliza este método se permite la utilización de los circuitos de un enlace por parte del tráfico alternativo hasta llegar a un umbral de $c-r$, por encima del cual los circuitos disponibles sólo podrán ser utilizados por el tráfico directo del enlace. Aparece ahora un parámetro nuevo: la reserva, r . El desarrollo es similar al mostrado en 2.2, usando también el modelo EFP.

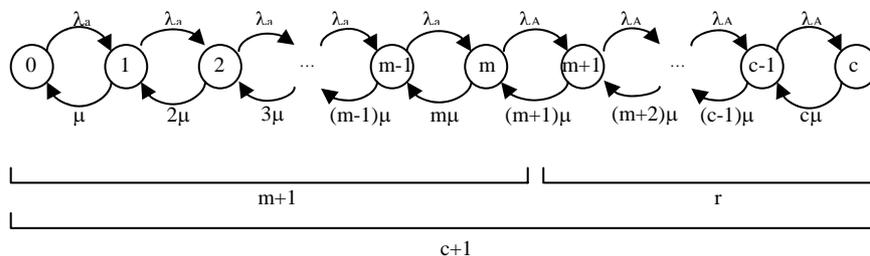


Figura 2.12 – Diagrama de estados de un enlace con reserva de troncales.

Como se muestra en el diagrama de estados de la figura 2.12, el tráfico ofrecido al enlace es distinto dependiendo de si el estado respeta o no la reserva. Así, la tasa de llamadas ofrecida al enlace es

$$\lambda_o(n) = \begin{cases} \lambda_a & 0 \leq n < m \\ \lambda_A & m \leq n < c \end{cases}$$

$$\lambda_a = \lambda_{directo} + \lambda_{alternativo} \quad (2.24)$$

$$\lambda_A = \lambda_{directo}$$

Para calcular la expresión de la probabilidad de estado se plantean las ecuaciones locales de balance de flujo en un estado genérico n . En particular, cuando $0 \leq n \leq m$ se tiene que

$$\lambda_a p(n-1) = n\mu p(n) \quad (2.25)$$

que, mediante recursividad, origina

$$p(n) = \frac{(\lambda_a/\mu)^n}{n!} p(0) = \frac{a_o^n}{n!} p(0) \quad (2.26)$$

De forma análoga, para $m < n \leq c$ se tiene que

$$p(n) = \frac{(\lambda_A/\mu)^{n-m}}{n(n-1) \cdots (m+1)} p(m) = \frac{A_o^{n-m}}{n(n-1) \cdots (m+1)} p(m) \quad (2.27)$$

con A_o y a_o los tráficos directo y total, respectivamente, ofrecidos al enlace. El uso de la notación A_o para el tráfico ofrecido al enlace cuando éste se halla en zona de reserva, tráfico ofrecido *interior*, es posible ya que coincide precisamente con el tráfico ofrecido entre el par de nodos conectados por el enlace, es decir, el tráfico ofrecido *exterior*. No hay, pues, contradicción con la notación empleada en el apartado 2.2.

Calculando $p(m)$ con (2.26) y sustituyendo en (2.27) se obtiene la expresión general de la probabilidad de estado

$$p(n) = \begin{cases} \frac{a_o^n}{n!} p(0) & 0 \leq n \leq m \\ \frac{A_o^{n-m} a_o^m}{n!} p(0) & m \leq n \leq c \end{cases} \quad (2.28)$$

Se tienen, por lo tanto, dos expresiones dependiendo del estado n , sin que se conozca la relación entre A_o y a_o . Sea un nuevo parámetro k la porción del tráfico total ofrecido al enlace correspondiente a llamadas reencaminadas, rechazadas por su ruta directa. El tráfico ofrecido directo se puede calcular como

$$A_o = a_o(1-k) \quad 0 \leq k \leq 1 \quad (2.29)$$

que se corresponde con el tráfico ofrecido total (a_o) menos el tráfico ofrecido alternativo ($a_o k$).

Reescribiendo (2.28) en función de k y sustituyendo en (2.29)

$$p(n) = \begin{cases} \frac{a_o^n}{n!} p(0) & 0 \leq n \leq m \\ \frac{a_o^n (1-k)^{n-m}}{n!} p(0) & m \leq n \leq c \end{cases} \quad (2.30)$$

con $p(0)$ calculable a partir de la suma de probabilidades de estado igual a la unidad:

$$p(0) = \left[\sum_{n=0}^m \frac{a_o^n}{n!} + \sum_{n=m+1}^c \frac{a_o^n (1-k)^{n-m}}{n!} \right]^{-1} \quad (2.31)$$

El cálculo de la *probabilidad de pérdida en llamadas* es distinto al considerar reserva de troncales, ya que un enlace se considerará bloqueado en lo que a tráfico alternativo se refiere, cuando se halle dentro de la zona de reserva para tráfico directo, no cuando esté realmente bloqueado. Así, la probabilidad de bloqueo de un enlace en tráfico alternativo es

$$p_m = \Pr\{n \geq m\} = p(0) \sum_{n=m}^c \frac{a_o^n (1-k)^{n-m}}{n!} = 1 - p(0) \sum_{n=0}^{m-1} \frac{a_o^n}{n!} = 1 - \overline{p}_m \quad (2.32)$$

$$\overline{p}_m = p(0) \sum_{n=0}^{m-1} \frac{a_o^n}{n!} \quad (2.33)$$

por lo que la *probabilidad de pérdida en llamadas*, obtenida en la ecuación (2.9), queda

$$P = p \left[1 - (1 - p_m)^2 \right]^M = p \left[1 - \overline{p}_m^2 \right]^M \quad (2.34)$$

siendo p la probabilidad de bloqueo de un enlace en tráfico directo

$$p = p(c) = p(0) \frac{a_o^c}{c!} (1-k)^r \quad (2.35)$$

con $r=c-m$.

Queda, ahora, calcular la relación entre los tráficos ofrecidos directo y total. El tráfico total cursado por un enlace es la suma de los tráficos cursados directo y alternativo:

$$a_c = a_{c,dir} + a_{c,alt} \quad (2.36)$$

donde el tráfico cursado directo, teniendo en cuenta que A_o es el tráfico ofrecido directo, se puede expresar

$$a_{c,dir} = A_o (1-p) \quad (2.37)$$

mientras que el tráfico cursado alternativo, teniendo en cuenta que $a_{o,alt} = a_o k$, según vimos en la expresión (2.29), puede expresarse de la forma

$$a_{c,alt} = a_{o,alt} \Pr\{n < m\} = a_o k \overline{p}_m \quad (2.38)$$

y también de la forma

$$a_{c,alt} = 2A_o p \left[1 - (1 - \overline{p_m^2})^M \right] = 2A_o p \left(1 - \frac{P}{p} \right) = 2A_o (p - P) \quad (2.39)$$

si tenemos en cuenta el resultado obtenido en la expresión (2.6). El dígito 2 representa el hecho de que los enlaces son bidireccionales. La probabilidad de uso de un determinado nodo como tándem entre otros dos ($1/M$) se simplifica al multiplicar la expresión por el número de posibles rutas alternativas (M) que pueden atravesar un enlace determinado en un sentido.

Para relacionar A_o y a_o pueden igualarse (2.38) y (2.39), obteniendo

$$a_o = A_o \frac{2(p - P)}{k \overline{p_m}} \quad (2.40)$$

También puede obtenerse una expresión similar sustituyendo (2.37) y (2.38) en (2.36), y (2.37) y (2.39) en (2.36)

$$a_c = A_o(1 - p) + a_o k \overline{p_m} = a_o(1 - k)(1 - p) + a_o k \overline{p_m} \quad (2.41)$$

$$a_c = A_o(1 - p) + 2A_o(p - P) = A_o(1 - p - 2P) \quad (2.42)$$

e igualando las dos expresiones del tráfico cursado total para obtener

$$a_o = A_o \frac{1 + p - 2P}{(1 - k)(1 - p) + k \overline{p_m}} \quad (2.43)$$

2.6.2. Resolución iterativa

El desarrollo teórico cuando no se emplea reserva de troncales ha conducido, en apartados anteriores, a una ecuación explícita no resoluble de forma analítica, pero sí mediante un método iterativo.

En esta ocasión, no se obtiene una ecuación explícita con una variable despejada, por lo que no es posible el cálculo de raíces, ni siquiera de forma iterativa. Por el contrario, para cada valor del parámetro conocido (el tráfico ofrecido entre cada par de nodos, A_o) se obtiene una serie de ecuaciones interdependientes en las que se desconoce el valor de las variables involucradas. Por ello, se ha de emplear un procedimiento que asigne un valor inicial a alguna de las variables y emplee las expresiones antes obtenidas para calcular las restantes. La exactitud de estos resultados es fuertemente dependiente del valor inicial dado a la variable, por lo que ha de iterarse el procedimiento hasta lograr la resolución deseada. El principal problema que se presenta ahora es el desconocimiento de la convergencia del método o, caso de existir varias soluciones, de la solución hacia la que converge.

Una posible secuencia de iteración es la siguiente:

1. Inicializar a_o como A_o ; por lo tanto, por la definición (2.29), $k=0$.
2. Calcular $p(0)$ con (2.31).
3. Calcular $\overline{p_m}$ mediante (2.33).
4. Calcular la probabilidad de bloqueo de enlace, p , con (2.35).
5. Calcular la probabilidad de bloqueo de ruta, P , con (2.34), para lo que necesito p y $\overline{p_m}$.
6. Recalcular el tráfico ofrecido al enlace, a_o , mediante (2.43) y k por (2.29).

- Iterar los pasos 2 a 6 hasta que la diferencia entre los valores de a_o en dos iteraciones sucesivas sea inferior a un determinado umbral.

De esta forma conseguimos quedarnos con el valor de P , para el cual, se logra la convergencia del método. El hecho de recalcular a_o mediante (2.43) y no mediante (2.40) se debe a que la primera admite la posibilidad de que $k=0$, caso que siempre se da en la primera iteración por la inicialización de a_o .

2.6.3. Resultados

En este apartado se muestran las gráficas que se obtienen al resolver las ecuaciones antes planteadas para uno de los casos de inestabilidad estudiados en este mismo capítulo. Se verá que la reserva de un número pequeño de troncales ya soluciona el problema de la inestabilidad.

El caso sobre el que se trabajará es el de una red mallada de enlaces de 100 troncales en la que cada nodo dispone de tres rutas alternativas, además de la directa, para alcanzar otro nodo de la red.

Los resultados obtenidos son los que se muestran en la figura 2.14. Se han representado los mismo cinco parámetros que en apartados anteriores. En cada una de las gráficas se incluyen curvas para diferentes valores de troncales reservadas. Hay, sin embargo, dos casos particulares: $r=0$, que es la red sin control –el caso estudiado en 2.4.2– que ha probado ser inestable, y $r=100$, caso que coincide con el encaminamiento directo por hallarse todas las troncales del enlace reservadas.

Sin embargo, el verdadero efecto de la reserva de troncales se aprecia en la diferencia entre las curvas con $r=0$ y $r=1$. Tan sólo con una troncal reservada para el tráfico directo se logra eliminar las inestabilidades, es decir, ya no hay ningún tráfico ofrecido para el que existen varias soluciones de cualquiera de los otros parámetros de red. Si bien para alta carga las prestaciones mejoran con la reserva – el caso límite es el encaminamiento directo; recuérdese que sus prestaciones en alta carga son mejores que las de cualquier método de encaminamiento alternativo–, no sucede lo mismo para cargas intermedias. Por lo tanto, la reserva óptima es la menor que elimina las inestabilidades.

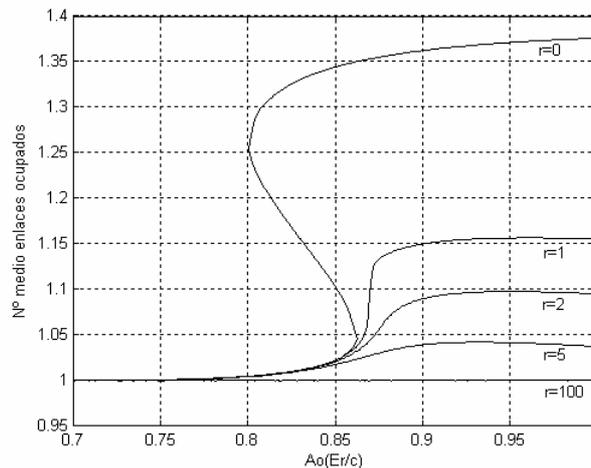


Figura 2.13 – Número medio de enlaces ocupados en una red simétrica totalmente conectada ($N=5$, $c=100$, $M=3$) frente al tráfico normalizado ofrecido entre cada par de nodos para diferente número de rutas reservadas.

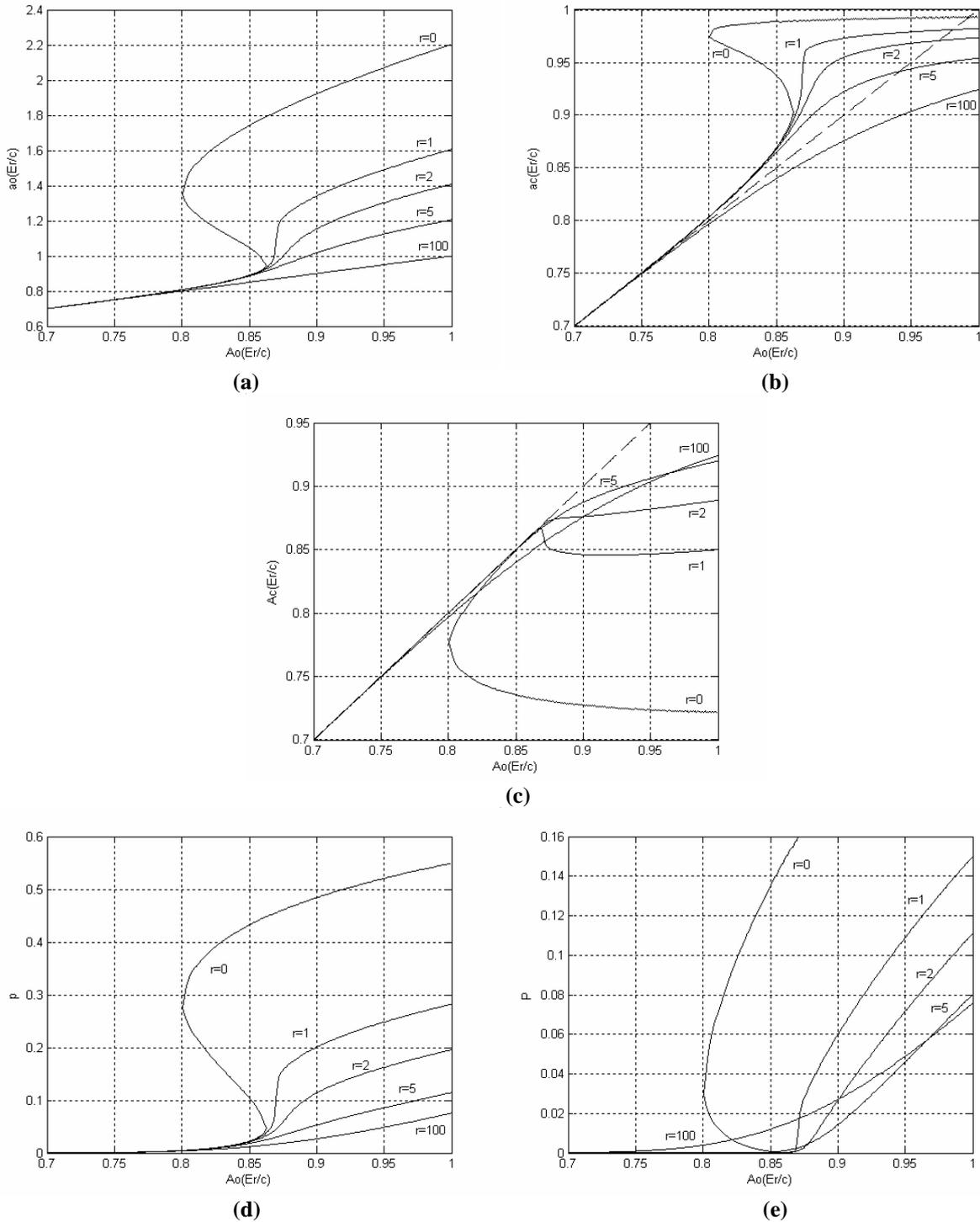


Figura 2.14 – Comportamiento de una red simétrica totalmente conectada ($N=5$, $c=100$, $M=3$) frente al tráfico normalizado ofrecido entre cada par de nodos para diferentes valores de reserva. **(a)** Tráfico normalizado ofrecido a cada enlace. **(b)** Tráfico normalizado cursado por cada enlace. **(c)** Tráfico normalizado cursado por la red entre cada par de nodos. **(d)** Probabilidad de bloqueo de un enlace. **(e)** Probabilidad de bloqueo de ruta.

En la figura 2.13 se ha representado el efecto de la reserva de troncales sobre el número medio de enlaces ocupados por las llamadas en curso. Además de eliminar la inestabilidad, el hecho de reservar un número relativamente bajo de troncales para tráfico directo –tráfico que usa un solo enlace– reduce notablemente el número de llamadas que emplean rutas alternativas. Nótese una vez más que niveles de reservas alto tienden al encaminamiento exclusivo por la ruta directa.

Se refuerza una vez más que la causa de la inestabilidad sea un sobreuso de enlaces por las llamadas encaminadas a través de una ruta alternativa, ya que la limitación de recursos para tráfico alternativo elimina tal deficiencia.

2.7. Métodos de encaminamiento dinámico

2.7.1. Breve historia del encaminamiento dinámico

El desarrollo del encaminamiento alternativo dinámico en redes de conmutación de circuitos ha conocido varias etapas:

- **Génesis (1975-80):** La posibilidad de realizar un control flexible del tráfico de la red dio lugar a la formulación de varios problemas teóricos. Entre ellos, el encaminamiento de tráfico en tiempo real ha sido comúnmente reconocido como el más prometedor. Los primeros trabajos estaban relacionados con los conceptos de control de tráfico para redes con encaminamiento alternativo. Una de las primeras propuestas teóricas para el encaminamiento dinámico la dio, a mediados de los 70, K. Narendra cuando introdujo el uso de técnicas de aprendizaje o *learning automata* para el encaminamiento del tráfico telefónico. Un seminario impartido por Narendra en los Laboratorios Bell permitió concentrar los esfuerzos de AT&T para estudiar, y finalmente implementar una red DNHR (*Dynamic NonHierarchical Routing*). Este trabajo vino acompañado por varios estudios que dieron lugar a nuevos métodos de encaminamiento dinámico.
- **Época de Expansión (1979-91):** El despegue del encaminamiento dinámico ocurrió a principios de los 80, a partir de que Bell Northern Research (BNR), ahora parte del Nortel Networks, estableciera un campo de trabajo en Toronto en 1979, basado en la realización periódica de medidas del estado de la red. Al mismo tiempo, varios centros de investigación de los operadores de red, así como varias universidades, establecieron otras propuestas. Todo ello dio lugar al desarrollo a gran escala del DNHR en las redes de larga distancia de AT&T en 1984, seguido del despliegue del campo de trabajo STAR en la red metropolitana de París por parte de France Telecom en 1987. De la misma forma, también se implantó un sistema de encaminamiento dinámico local en los conmutadores de Italcable para el acceso a la red internacional. En ese tiempo, el encaminamiento dinámico fue reconocido como un tema fundamental por la comunidad de investigación de telecomunicaciones. De hecho, se introdujeron nuevas propuestas en los congresos internacionales de teletráfico, que después fueron discutidas por los organismos de estandarización como el CCITT (ahora ITU-T). En esta época se hizo una gran divulgación del encaminamiento dinámico a través de varios seminarios, las publicaciones del *IEEE* (que proporcionaron un estudio comprensible de las técnicas de encaminamiento dinámico), así como la publicación de recomendaciones ITU-T (que recalcan la importancia del encaminamiento dinámico).
- **Época de Madurez (1991-2000):** Durante esta década, muchos operadores implementaron sistemas de encaminamiento dinámico en sus redes domésticas y en las redes internacionales. En particular, el sistema WIN (*Worldwide Intelligent Network-Routing*), que se inició a principio de los 90, es una iniciativa que adoptaron varios operadores de red internacionales. Las sucesivas publicaciones del *IEEE* enfocaron las principales implementaciones, y varias recomendaciones ITU-T complementaron los estándares existentes.
- **Época de Recesión (desde 2000):** Con la aparición de nuevas tecnologías como ATM e IP, que son orientadas a conmutación de paquetes, la preocupación por las redes TDM ha decrecido rápidamente, y los nuevos trabajos aparecidos se preocupan por principios similares, de QoS y de encaminamiento, aplicados a otros tipos de redes. Se espera que con el desarrollo de las redes IP los actuales sistemas de encaminamiento dinámico desaparezcan progresivamente, y sean reemplazados por otros nuevos sistemas implementados en protocolos IP.

2.7.2. Clasificación de los métodos de encaminamiento alternativo

Un método específico de encaminamiento alternativo se caracteriza por la *tabla de encaminamiento* usada en el método. La tabla de encaminamiento posee una serie de *reglas* para seleccionar un camino para una petición de conexión determinada. Así, cuando una llamada llega a su conmutador origen, que implementa un método de encaminamiento alternativo, ejecuta las *reglas de selección* de ruta asociadas con la tabla de encaminamiento para determinar una ruta. Además, en un determinado método de encaminamiento, el conjunto de rutas que pueden ser asignadas a la llamada se puede alterar de acuerdo a una cierta *regla de alteración*. Por otro lado, dependiendo de si un conmutador funciona como origen, destino, o *tándem*, se usan diferentes tablas de encaminamiento para la llamada. Normalmente, es el conmutador origen el que determina el método de encaminamiento alternativo usado para la llamada y selecciona la tabla de encaminamiento apropiada.

Una red se puede controlar con un control de llamada progresivo, un control de llamada en origen, o una mezcla de los dos métodos de control. En una red con control de llamada progresivo cada conmutador debe seleccionar los enlaces a utilizar hacia el siguiente conmutador. En una red con control de llamada en origen es el conmutador origen el que mantiene todo el control de la llamada.

Las rutas en una tabla de encaminamiento pueden tomar varias formas. Pueden consistir en un grupo de enlaces directos, en rutas con dos grupos de enlaces a través de un solo nodo *tándem*, o en rutas con múltiples grupos de enlaces a través de múltiples nodos *tándem*. Para prevenir inestabilidades de la red o establecer prioridades entre diferentes tipos de llamadas, las rutas pueden estar sujetas a restricciones SCR (*Selective Circuit Reservation*). SCR requiere que algo más que el número específico de circuitos necesarios estén libres en cada grupo de enlaces antes de que una conexión se establezca.

En las recomendaciones ITU-T E.170, E.177 y E.DYN, los métodos de encaminamiento alternativo se clasifican en los siguientes cuatro tipos: FR (*Fixed Routing*), TDR (*Time-Dependent Routing*), SDR (*State-Dependent Routing*) y EDR (*Event-Dependent Routing*). Los métodos TDR, SDR y EDR pertenecen al encaminamiento alternativo dinámico; los métodos FR, como su propio nombre indica, lo hacen al encaminamiento alternativo fijo. Vamos a comentar cada uno de estos métodos en los siguientes párrafos.

- FR

En un método FR, las tablas de encaminamiento son fijas para cada llamada. Este tipo de métodos se puede utilizar tanto para encaminamiento jerárquico como no jerárquico. El conjunto de rutas posibles y la secuencia de selección de rutas se encuentran preestablecidas y se mantienen durante un largo periodo de tiempo. FR es el método de encaminamiento que debe utilizarse en una red jerárquica.

-TDR

Los métodos TDR son un tipo de encaminamiento alternativo dinámico en los que las tablas de encaminamiento se modifican en un instante de tiempo fijo del día o de la semana. Las tablas de encaminamiento TDR están preestablecidas y se mantienen durante un periodo de tiempo. Normalmente se establecen considerando la variación de la carga de tráfico en la red a lo largo del tiempo y se coordinan para evitar que coincidan en las horas cargadas de tráfico. DNHR es un ejemplo de método TDR.

Las tablas de encaminamiento establecen las rutas óptimas para un número muy grande de posibles alternativas, con el objetivo de minimizar el coste de la red. Estas tablas son cargadas y almacenadas en todos los conmutadores de la red, y periódicamente son recalculadas y actualizadas. De esta forma, un conmutador no necesita información adicional una vez que se han cargado las tablas de encaminamiento. Esto está en contraste con el diseño de las tablas de encaminamiento en tiempo real, que se da en los métodos SDR y EDR.

Las reglas de selección de ruta utilizadas en las tablas de encaminamiento TDR pueden consistir en un simple encaminamiento secuencial. En el método secuencial todo el tráfico en un periodo de tiempo dado se ofrece a un único conjunto de rutas. El tráfico desbordado en la primera ruta se reencamina a la segunda ruta, y así sucesivamente. Otras reglas de selección de ruta TDR pueden emplear técnicas probabilísticas para seleccionar cada camino.

Los caminos en la tabla de encaminamiento están sujetos a las restricciones de profundidad de búsqueda (*Depth-of-Search, DoS*). DoS requiere que la cantidad de ancho de banda disponible en cada enlace del camino sea suficiente para alcanzar un determinado umbral, que se difunde por la red a través de un mensaje de establecimiento. Las restricciones DoS previenen de que las conexiones que se encaminan a través del camino directo, por ejemplo, se vean inundadas por el encaminamiento de conexiones alternativas.

A continuación se da un ejemplo de establecimiento de conexión TDR. El primer paso para el conmutador es identificar el destino, y la ruta de la tabla de encaminamiento asociada a ese destino. El conmutador origen comprueba la capacidad disponible en el primer camino, y para ello proporciona a todos los nodos *tándem* los parámetros DoS. Cada conmutador *tándem* compara el ancho de banda disponible en cada enlace del camino con el umbral DoS. Si hay suficiente capacidad, el conmutador *tándem* correspondiente encamina la petición de establecimiento de conexión al siguiente conmutador, que ejecuta un procedimiento similar. Si la capacidad es insuficiente, el conmutador *tándem* envía un mensaje con un parámetro de ancho de banda no disponible hacia el conmutador origen, que deberá intentar el siguiente camino establecido por las reglas de selección de ruta de la tabla de encaminamiento.

- SDR

En los métodos SDR las tablas de encaminamiento se alteran automáticamente en función del estado de la red. Para un método SDR dado, las reglas de selección de ruta de la tabla de encaminamiento se implementan con el objetivo de modificar las elecciones de ruta en respuesta al cambio de estado de la red, y utilizan el principio de encaminar las llamadas por la mejor ruta disponible basándose en la información del estado de la red. Por ejemplo, en el método LLR (*Least Loaded Routing*), se calcula la capacidad residual de las rutas, y se selecciona la ruta con una mayor capacidad residual. Así, en general, los métodos SDR calculan el coste de la ruta para cada llamada, basándose en factores como el estado de la carga o el estado de congestión en los grupos de enlace de la red. Las rutas elegidas se usan durante un periodo de tiempo relativamente corto. La información del estado de la red se puede conseguir periódicamente o bajo demanda. DCR (*Dynamically Controlled Routing*), WIN (*Worldwide Intelligent Network routing*) y RTNR (*Real-Time Network Routing*) son ejemplos de métodos SDR.

En SDR, las tablas de encaminamiento son diseñadas por el conmutador origen o por un procesador de encaminamiento central con la ayuda de la información de red obtenida de los conmutadores de la red. Hay varias implementaciones de SDR, que se diferencian en

- a) Si el cálculo de las tablas de encaminamiento es distribuido entre los conmutadores de la red o es centralizado, y realizado por un procesador de encaminamiento.
- b) Si el cálculo de las tablas de encaminamiento se realiza periódicamente o llamada a llamada.

Esto da lugar a tres implementaciones diferentes de métodos SDR:

- a) SDR centralizado periódico: El procesador de encaminamiento centralizado obtiene periódicamente de los distintos conmutadores el estado de los grupos de enlaces e información del estado del tráfico, y realiza el cálculo de la tabla de encaminamiento óptima periódicamente. Para establecer la tabla de encaminamiento, el procesador de encaminamiento ejecuta un procedimiento particular de optimización de la tabla de encaminamiento como LLR y transmite las tablas de encaminamiento a la red de conmutadores. DCR es un ejemplo de SDR centralizado periódico.
- b) SDR distribuido periódico: Cada conmutador de la red obtiene periódicamente de los distintos conmutadores el estado de los grupos de enlaces e información del estado del tráfico, y realiza el

cálculo de la tabla de encaminamiento óptima periódicamente. Para establecer la tabla de encaminamiento, el conmutador origen ejecuta un procedimiento particular de optimización de la tabla de encaminamiento como LLR. WIN es un ejemplo de SDR distribuido periódico.

- c) SDR distribuido llamada a llamada: El conmutador origen de la red obtiene el estado de los grupos de enlaces e información del estado del tráfico del conmutador destino, y quizás del conmutador *tándem*, para cada llamada y ejecutan el cálculo de la tabla de encaminamiento óptima para cada llamada. Para determinar la tabla de encaminamiento óptima, el conmutador origen ejecuta un procedimiento particular de optimización como LLR. RTNR es un ejemplo de SDR distribuido llamada a llamada.

Los caminos en la tabla de encaminamiento están sujetos a las restricciones DoS en cada enlace, y los mecanismos de establecimiento de conexión son similares al ejemplo dado en el método TDR.

A continuación se muestran ejemplos de intercambio de información para el diseño de la tabla de encaminamiento para los tres casos de SDR. En el SDR centralizado periódico, por ejemplo, cada conmutador periódicamente (p.e. cada 10 segundos) envía información de estado hacia el procesador de encaminamiento, en la que se incluye información sobre el estado del tráfico y la carga en la red. En respuesta, el procesador de encaminamiento envía las recomendaciones de encaminamiento a cada conmutador periódicamente (p.e. cada 10 segundos), que contienen información sobre la ruta alternativa para cada par de conmutadores origen-destino. En el SDR distribuido periódico, cada conmutador periódicamente envía información de estado hacia cada conmutador de la red, en la que se comunica la información de estado de la red. En el SDR distribuido llamada a llamada, en primer lugar el conmutador origen intenta establecer una conexión por el camino directo, si falla, entonces este conmutador envía al conmutador destino una petición de estado. El conmutador destino responde al origen con la información de estado.

Por lo tanto, podemos establecer que en el intercambio de información para establecer la tabla de encaminamiento intervienen los siguientes mensajes:

- Petición de estado (QUERY): Se envía desde el conmutador origen hacia el conmutador destino, o desde el conmutador origen al procesador de encaminamiento, y se puede referir tanto al estado del conmutador como de los enlaces.
- Estado (STATUS): Se envía desde los conmutadores origen, destino o *tándem* hacia el procesador de encaminamiento, o desde el conmutador destino hacia el origen. También se puede referir tanto al estado del conmutador como de los enlaces.
- Información de recomendación (RECOMM): Envía las recomendaciones desde el procesador de encaminamiento hacia los conmutadores origen, destino o *tándem*.

- EDR

En los métodos EDR las tablas de encaminamiento se actualizan en caso de que la llamada falle o tenga éxito en una ruta determinada. En EDR una llamada se ofrece en primer lugar a una ruta fija y preestablecida, que normalmente coincide con la ruta directa. Si no se dispone de ningún circuito libre en la ruta preestablecida el tráfico desbordado se ofrece a una ruta alternativa seleccionada. Si la llamada se bloquea en la ruta alternativa actual, se selecciona otra ruta alternativa del conjunto de rutas posibles, de acuerdo a unas reglas de selección de ruta específicas. Por ejemplo, la elección de la ruta alternativa se puede realizar aleatoriamente, cíclicamente, o de otras formas, y la ruta seleccionada se puede mantener mientras que las llamadas se establezcan con éxito en esa ruta. MIR (*Multiple Intelligent Routing*), DAR (*Dynamic Alternative Routing*), STR (*State and Time-dependent Routing*) y LAW son ejemplos de métodos EDR.

En EDR las tablas de encaminamiento son diseñadas por el conmutador origen usando información de la red obtenida durante el establecimiento de la llamada, y posteriormente no se necesita más información. Típicamente, el conmutador origen primero selecciona el grupo de enlaces directos y si está ocupado se intenta la ruta alternativa que gana.

Los caminos en la tabla de encaminamiento están sujetos a las restricciones DoS en cada enlace, y los mecanismos de establecimiento de conexión son similares al ejemplo dado en el método TDR.

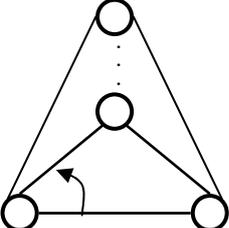
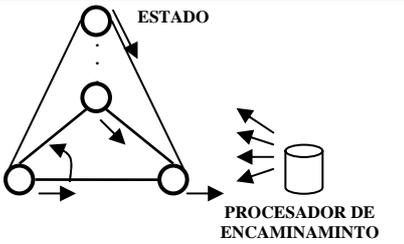
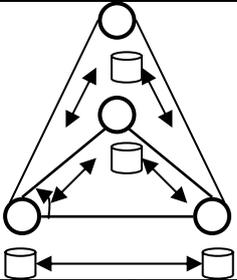
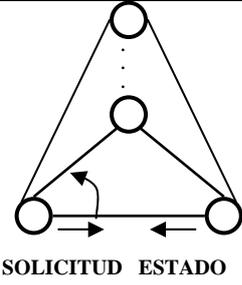
<p>Event Dependent Routing (por ejemplo: AMI, DAR, LAW/ODR)</p>		<ul style="list-style-type: none"> - Seleccionar el camino primario. - Cuando la llamada se bloquea, elegir aleatoriamente el nuevo camino alternativo. - Opcionalmente, permitir N intentos posibles para encontrar el camino alternativo con éxito.
<p>State Dependent Routing (centralizado, periódico) (por ejemplo: DCR)</p>		<ul style="list-style-type: none"> - Seleccionar el camino primario. - Cada nodo envía periódicamente el estado del enlace al procesador de encaminamiento. - Cada nodo recibe periódicamente del procesador de encaminamiento las actualizaciones.
<p>State Dependent Routing (distribuido, periódico) (por ejemplo: WIN)</p>		<ul style="list-style-type: none"> - Intercambio de tráfico de forma periódica entre todos los conmutadores de la red. - Seleccionar el camino primario.
<p>State Dependent Routing (distribuido, llamada a llamada) (por ejemplo: RTNR)</p>		<ul style="list-style-type: none"> - Seleccionar el camino primario. - Los nodos origen solicitan a los nodos destino el estado de los enlaces. - El encaminamiento cambia llamada a llamada.

Figura 2.15 – Implementaciones de métodos de encaminamiento dinámico.

2.7.3. Implementaciones actuales de encaminamiento dinámico

Hay muchos operadores que han implementado métodos de encaminamiento dinámico en sus redes domésticas e internacionales. En la tabla 2.2 se muestran las implementaciones actuales de varios sistemas de encaminamiento dinámico que se han desarrollado en varios tipos de redes. Como se puede ver, muchos permanecen en operación.

De forma adicional a las implementaciones operacionales, se ha realizado una gran cantidad de investigaciones a través de campos de trabajo, o estudio de introducción, o incluso a través de academias de investigación. La enorme bibliografía que existe de este tema refleja el enorme impacto que el encaminamiento dinámico tuvo en la comunidad de las telecomunicaciones durante la última década.

2.7.4. Otras clasificaciones

Existen, sin embargo, otras clasificaciones para los algoritmos de encaminamiento alternativo. La clasificación que vamos a seguir en el desarrollo de este proyecto, se basa en el método que se utiliza en

las reglas de selección de la ruta alternativa a utilizar para una llamada concreta. De esta forma, clasificamos los algoritmos de encaminamiento alternativo en:

- Encaminamiento alternativo fijo. Una nueva llamada intentará establecer una conexión por una ruta predefinida; si la ruta está bloqueada, se intentará por la ruta alternativa, también predefinida; si esta también está bloqueada la llamada es rechazada. No se intentan más rutas.
- Encaminamiento alternativo dinámico (rutas alternativas de dos enlaces). Si la primera ruta está bloqueada (que normalmente es una ruta directa de un enlace), se intentarán otras rutas alternativas. En la mayoría de los casos, sólo se consideran rutas alternativas de dos enlaces. Hay muchos esquemas para determinar cómo elegir la ruta alternativa. A continuación se muestran algunos de estos esquemas.
 - Encaminamiento aleatorio. Este tipo de encaminamiento elige aleatoriamente una ruta alternativa; si la ruta elegida está bloqueada, la llamada es rechazada. No se intentan otras rutas alternativas. DAR y STR son dos ejemplos de este tipo de encaminamiento.
 - Encaminamiento secuencial. Examina un subconjunto de rutas alternativas en una secuencia predefinida; si todas las rutas alternativas están bloqueadas, la llamada es rechazada.
 - Encaminamiento por ruta menos cargada. Elige la ruta alternativa con la mayor capacidad disponible; si todas las rutas alternativas están bloqueadas, la llamada es rechazada. Este método intenta distribuir la misma capacidad entre todas las rutas alternativas. LLR, RTNR y DCR son ejemplos de este tipo de encaminamiento.

En los capítulos 3 y 4 se van a estudiar en profundidad un ejemplo de encaminamiento aleatorio, y otro de encaminamiento por ruta menos cargada. Concretamente se van a mostrar las prestaciones del DAR (*Dynamic Alternative Routing*) y del LLR (*Least Loaded Routing*), respectivamente.

Podemos observar que existe una cierta relación entre las dos clasificaciones de métodos de encaminamiento alternativo presentadas en este capítulo. Algunas implementaciones de métodos de encaminamiento por ruta menos cargada, como RTNR y DCR, son a la vez métodos de encaminamiento SDR (*State Dependent Routing*); por su parte, el método de encaminamiento aleatorio DAR, también es un método de encaminamiento EDR (*Event Dependent Routing*).

Tipo de encaminamiento	Sistemas de encaminamiento dinámico	Red	Año de inicio	Año de finalización	Comentarios	
TDR	DNHR <i>Dynamic Non-Hierarchical Routing</i>	AT&T US National Network	1984	1991	DNHR fue sustituido por RTNR en 1991	
		AT&T FTS-2000 Network	1987	2002	DNHR fue sustituido por RTNR en 2002	
SDR <i>centralizado periódico</i>	GTAI (GTAI es un acrónimo italiano: <i>Management of the Traffic Transit Italcable switches</i>)	Italcable	1984	1985	Este mecanismo de encaminamiento fue implementado entre los 3 conmutadores internacionales operados por Italcable	
		DCR <i>Dynamically Controlled Routing</i>	Stentor Canada National Network	1991	En operación	Conocido como HPR (<i>High Performance Routing</i>)
			Bell Canada Network	1992	En operación	Consiste en una red local DCR en el área de Toronto y otra en el área de Montreal
			Sprint National Network	1994	En operación	
			MCI US National Network	1995	En operación	
Qwest Communications National Network	1999	En operación				
SDR <i>distribuido periódico</i>	WIN <i>Worldwide Intelligent Network routing</i>	Worldwide Intelligent Network	1993	En operación		
SDR <i>distribuido llamada a llamada</i>	RTNR <i>Real-Time Network Routing</i>	AT&T US National Network	1991	En operación		
		AT&T FTS-2000 Network	2002	En operación		
	RINR <i>Real-Time Internetwork Routing</i>	AT&T Global International Network	1991	En operación		
EDR	STR <i>State and Time-dependent Routing</i>	NTT Japan National Network	1992	2002		
	DAR <i>Dynamic Alternative Routing</i>	British Telecom UK National Network	1993	?		
	STT <i>Success-To-the-Top network routing</i>	AT&T US National Network	1995	1999	STT es un método usado para encaminar llamadas durante un periodo de tiempo con dispositivos de mejora de la voz en el camino	
	LAW <i>Lastabhängige Automatische Wegesuche</i> (en inglés, Automatic Last Choice Routing)	Deutsche Telecom National Network	1995	En operación	LAW se implementó en tanto en las redes de tránsito como en las redes de acceso nacionales e internacionales	
	AMI <i>Acheminement Multiple Intelligent</i> (en inglés, Multiple Intelligent Routing – MIR)	France Telecom Long Distance Network	1998	En operación	AMI/MIR es un sistema EDR con múltiples rutas de desbordamiento.	

Tabla 2.2 – Estado de las redes de encaminamiento dinámico.

Capítulo 3

Encaminamiento aleatorio.

DAR

3.1. Introducción

El encaminamiento dinámico en redes de pérdidas pretende disminuir el GoS de la red permitiendo el uso de rutas alternativas por donde cursar una llamada. Una de las razones que origina la ganancia obtenida al usar encaminamiento dinámico es la no uniforme distribución del tráfico en todos los enlaces de la red, por el propio carácter aleatorio de las llamadas, por la no coincidencia de la hora cargada en toda la red, etc. Así, puede haber zonas de la red saturadas y zonas con recursos libres. El algoritmo de encaminamiento debería de detectar y solventar estas situaciones, encaminando por las zonas con recursos libres.

En definitiva, lo que se busca es que el algoritmo de encaminamiento dinámico uniformice la distribución del tráfico en los nodos y enlaces que componen la red. La filosofía que se describe en este capítulo radica en la búsqueda de la uniformidad mediante la generación de una variable aleatoria discreta uniforme, cuyos posibles valores se asocian a la ruta por la que encaminar.

3.2. Descripción del DAR (Dynamic Alternative Routing)

El DAR es el algoritmo de encaminamiento empleado por British Telecom en su red nacional y en la red de acceso a los ISCs o centros internacionales de conmutación. Se trata de un algoritmo muy sencillo perteneciente a la familia del encaminamiento aleatorio. Surge como evolución del AAR, el anterior algoritmo empleado por British Telecom, que estaba basado en la búsqueda de la uniformización mediante la selección cíclica de una ruta entre todas las posibles.

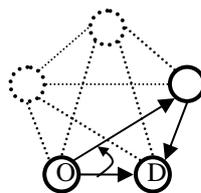


Figura 3.1 – Modo de operación del DAR.

El modo de operación del DAR se representa en la figura 3.1, y consiste en que cada par de nodos de la red posee un enlace directo que los une –se estudiará en una red simétrica totalmente conectada– y tienen asociado otro nodo de la red que actúa como nodo *tándem*. Cuando se produce una llamada, el algoritmo intenta cursarla por la ruta directa. Si el enlace directo está bloqueado, la llamada es encaminada hacia los dos enlaces que unen los nodos origen, *tándem* y destino. Si la ruta alternativa puede cursar la llamada, la siguiente llamada rechazada por el enlace directo también se intentará por esta misma ruta secundaria. Sólo cuando ésta también esté bloqueada la llamada será definitivamente rechazada, y un nuevo nodo, y por lo tanto una nueva ruta, será elegido aleatoriamente entre los restantes posibles como nodo y ruta alternativos para el par de nodos origen-destino de la llamada rechazada.

Existe, por lo tanto, una analogía con el algoritmo de búsqueda de la ruta alternativa de forma totalmente aleatoria. El primer intento de cursar la llamada siempre es a través del enlace directo entre los nodos que se desea conectar. Esto es lógico, ya que es la ruta que menos perjuicio causa a la red por el mínimo uso de recursos, como se ha visto en el capítulo anterior. Cualquier ruta alternativa emplea, como

mínimo, el doble de recursos, ya que consta de dos o más enlaces (en este estudio se supondrán compuestas por dos) más los recursos de conmutación en el nodo, que normalmente no son considerados a ese nivel de análisis.

De la misma forma, existe una diferencia con el algoritmo aleatorio puro. El encaminamiento por la ruta alternativa es fijo para cada par de nodos, residiendo la aleatoriedad en la elección de la ruta secundaria cuando la considerada hasta ese momento se bloquea. Este método tiene un efecto de *empaquetamiento de llamadas* sobre la misma ruta alternativa, de forma que el tráfico desbordado de un par de nodos puede no interferir en el de otro par. Esta política se conoce como *sticky-random* o *back to the winner*, refiriéndose, respectivamente, a que se mantiene la ruta aleatoriamente seleccionada o a que se encamina por la ruta alternativa que ha dado un resultado positivo más recientemente.

La peculiaridad anterior redundante en un menor coste computacional del algoritmo, característica que será estudiada de forma secundaria en el presente proyecto, pero de gran importancia de cara a la implementación del algoritmo en una red real. De hecho, hay algoritmos sólo implementables en una versión simplificada del mismo o no implementables, debido a que su complejidad es superior a las prestaciones en tiempo real del estado de la tecnología.

Otras dos características esenciales del DAR son su carácter distribuido y su desconocimiento del estado exacto de la red. Por una parte, el carácter distribuido implica que no existe un nodo central que gestione todas las tareas de encaminamiento; es más, cada nodo de la red toma las decisiones de encaminamiento independientemente de los otros nodos, basándose exclusivamente en el comportamiento de una ruta alternativa para cada nodo destino. Por otra parte, se logra hacer un encaminamiento dinámico y, por tanto, dependiente del estado de la red sin un conocimiento explícito del mismo, sino indirecto a través de la indicación de que una llamada no puede ser encaminada por una ruta alternativa particular, confesión implícita del bloqueo en alguno de los dos enlaces que la componen. Así, no es necesario que los nodos conozcan el estado de todos los recursos de la red, información que habrían de intercambiarse entre ellos mediante mensajes periódicos de señalización.

En el caso de decisiones centralizadas también habrían de existir mensajes de señalización desde el nodo central hacia los nodos de conmutación informando de la decisión tomada. De cara a los costes, la alternativa por la que opta el DAR evita, por una parte, los costes del nodo central, en el caso de que exista, y, mucho más importante, los costes de una red de señalización asociada. De cara a minimizar el tiempo de establecimiento de la conexión, tanto la implementación distribuida como el reducido coste computacional hacen al DAR enormemente atractivo. Añadiendo sus buenas prestaciones desde el punto de vista del GoS, se comprende, como se ha dicho al comienzo del capítulo, que esté implementado en una red real.

Con el fin de analizar de forma práctica las prestaciones del DAR, se ha realizado la programación de un simulador en lenguaje C que implementa dicho algoritmo de encaminamiento. A partir de aquí, todos los resultados mostrados han sido obtenidos mediante dicho simulador.

3.3. Prestaciones en redes multiclase

3.3.1. Distribución de los tiempos de servicio y entre llamadas

Interesa estudiar cuán robusto es el DAR ante variaciones en las distribuciones del tiempo entre llegadas y del tiempo de duración de la conexión. Esto dará una idea de la robustez del DAR ante la introducción de nuevos servicios multimedia.

Los resultados de la simulación aparecen reflejados en las figuras 3.2a y 3.2b, para variaciones del tiempo de duración del servicio y del tiempo entre llegadas, respectivamente. La distribución estadística del tiempo de conexión apenas influye, mientras que sí lo hace la del tiempo entre peticiones, mostrándose también mayores valores de GoS para tráfico de Poisson y menor para tiempo entre llamadas de duración fija, por la mayor varianza de la exponencial y varianza nula de la determinista.

La razón por la cual la distribución estadística del tiempo de conexión apenas influye, viene dada por el hecho de que en los sistemas sin colas el tiempo de servicio residual medio es independiente de la forma que tome la distribución del tiempo de servicio, tomando en cada caso el valor que le corresponda, que no tiene porqué coincidir con la media o cualquier otro parámetro de la distribución.

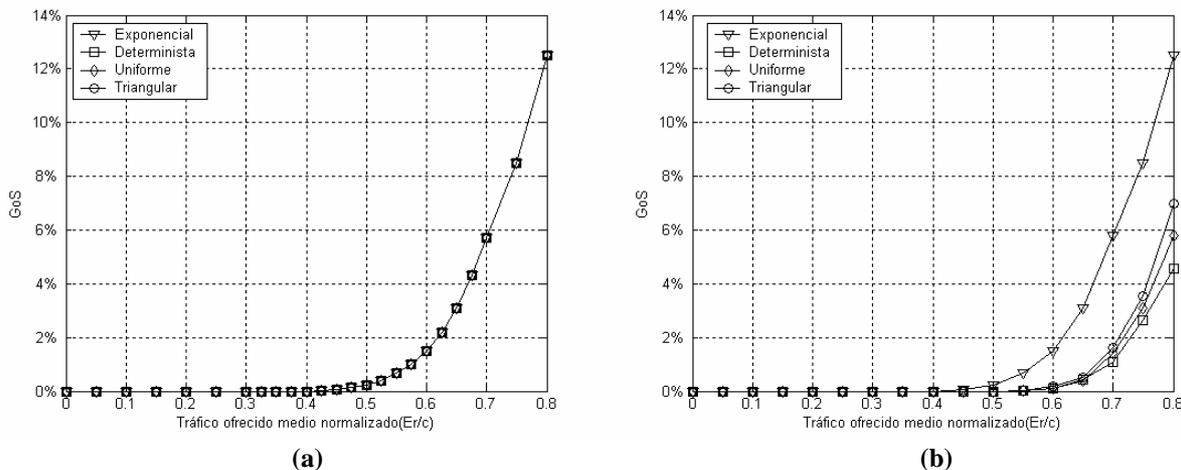


Figura 3.2 – Dependencia del GoS con la distribución del tráfico en una red totalmente conectada simétrica; $N=5$, $c=10$. **(a)** GoS para tiempo entre llamadas exponencial y tiempo de servicio variable. **(b)** GoS para tiempo de servicio exponencial y tiempo entre llamadas variable.

3.4. Comparación de prestaciones con el encaminamiento directo

En el estudio teórico del encaminamiento dinámico del capítulo anterior se obtuvieron algunos resultados que se pretenden comparar ahora con los resultados de las simulaciones realizadas.

3.4.1. Resultados generales

Son varios los parámetros que influyen en el GoS, sin tener en cuenta el algoritmo de encaminamiento. Entre ellos cabe destacar:

- Distribución estadística del tráfico: se supondrá tráfico de Poisson, es decir, con tiempos entre peticiones y de duración de servicio exponenciales.
- Topología de red, tanto en número de nodos de la misma, como en las capacidades y tráficos entre cada par de nodos: se supondrá una red simétrica para evitar la interferencia de las asimetrías en el GoS. Como en varias simulaciones anteriores, se ha elegido $N=5$ y $c=10$. En siguientes apartados se verá el efecto de estos dos parámetros.

Interesa conocer cuál es la mejora que el algoritmo considerado introduce, por lo que se comparan las prestaciones obtenidas con las que ofrece el encaminamiento directo. Son tres los parámetros que pueden permitir la comparación de prestaciones: GoS, tráfico cursado y capacidad de los enlaces, aunque sólo los dos primeros se usarán en este proyecto.

Para establecer la comparación se define el concepto de ganancia como la diferencia de prestaciones entre ambos algoritmos, referidas, o no, a las del directo, dependiendo de si es ganancia relativa o absoluta. Así, la ganancia relativa en un parámetro en función del otro se define como:

$$G_{GoS}^{DAR}(A_o) = \frac{GoS_{dir}(A_o) - GoS_{DAR}(A_o)}{GoS_{dir}(A_o)} = 1 - \frac{GoS_{DAR}(A_o)}{GoS_{dir}(A_o)} \tag{3.1a}$$

$$G_{A_c}^{DAR}(GoS) = \frac{A_{c,DAR}(GoS) - A_{c,dir}(GoS)}{A_{c,dir}(GoS)} = \frac{A_{c,DAR}(GoS)}{A_{c,dir}(GoS)} - 1 \quad (3.1b)$$

calculando el tráfico cursado a partir del ofrecido y del GoS

$$A_c = A_o(1 - GoS) \quad (3.2)$$

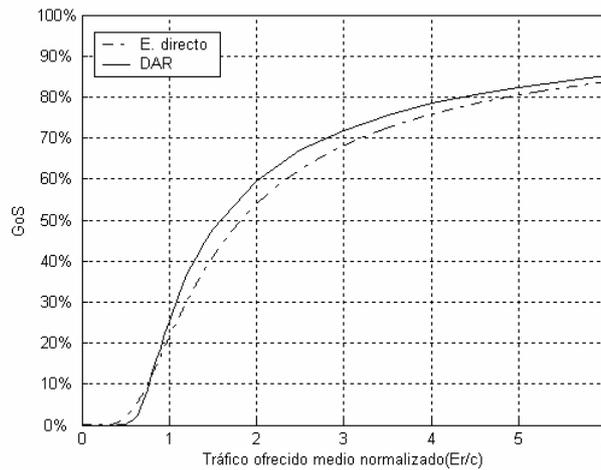


Figura 3.3 – Prestaciones del encaminamiento directo y del DAR en alta carga.

En la figura 3.3 se representan conjuntamente los GoSs que se obtienen con el encaminamiento directo y con el DAR para un amplio margen de tráfico. Se confirma por simulación lo que ya se predijo en el capítulo anterior: los algoritmos que contemplan el encaminamiento alternativo presentan mejores prestaciones para tráficos bajos, donde las rutas alternativas solventan el problema transitorio de bloqueo en el enlace directo, y peores para cargas altas, donde el uso de dos enlaces es más frecuente y provoca un aumento en el número de llamadas rechazadas. Recuérdese que uno de los parámetros característicos de los servicios es el GoS máximo en la hora cargada y que éste valor no suele superar el 2%.

Los resultados de la simulación, dentro del margen significativo de valores, se muestran en las figuras 3.4a a 3.4f, donde las figuras relativas a las prestaciones en GoS son las tres de la izquierda y las relativas a tráfico cursado las de la derecha. Las dos figuras superiores muestran las curvas de evolución de un parámetro frente al otro para encaminamiento directo y DAR, las centrales la ganancia absoluta y las inferiores la ganancia relativa, calculada según (3.1a) y (3.1b).

Debido a que sólo se representa dentro del margen más significativo de valores, se comprueban las mejores prestaciones del DAR. Sin embargo, las curvas de prestaciones (figuras 3.4a y 3.4b) se van acercando a medida que se aumenta el tráfico, anunciando la situación mostrada en la figura 3.3 y prevista con el modelo teórico. Ello también puede verse en cualquiera de las curvas de ganancia, tanto absoluta como relativa, donde se toma un valor siempre positivo, pero decreciente al aumentar el tráfico. Más allá del punto de cruce de las curvas de prestaciones de ambos algoritmos (observable en la figura 3.3), la ganancia tomará un valor negativo, indicando el empeoramiento de las prestaciones al usar rutas alternativas de dos enlaces.

Dado que ambas curvas de prestaciones (encaminamiento directo y DAR) son coincidentes para tráficos muy pequeños y se cruzan a partir de un determinado valor, la diferencia entre ambas presentará un máximo en el interior del intervalo definido. Los valores máximos de ganancia y los puntos en los que se producen se muestran de forma aproximada en la tabla 3.1.

A partir de las figuras 3.4a y 3.4b resultan evidentes los valores de ganancia absoluta, pues éstos representan la simple diferencia en prestaciones. Sin embargo, las ganancias relativas mostradas en las figuras 3.4e y 3.4f muestran que los máximos se obtienen para valores muy bajos de tráfico ofrecido o GoS. Esto es debido a la normalización que se realiza dividiendo por las prestaciones del encaminamiento directo, valores menores para menores tráficos o GoSs.

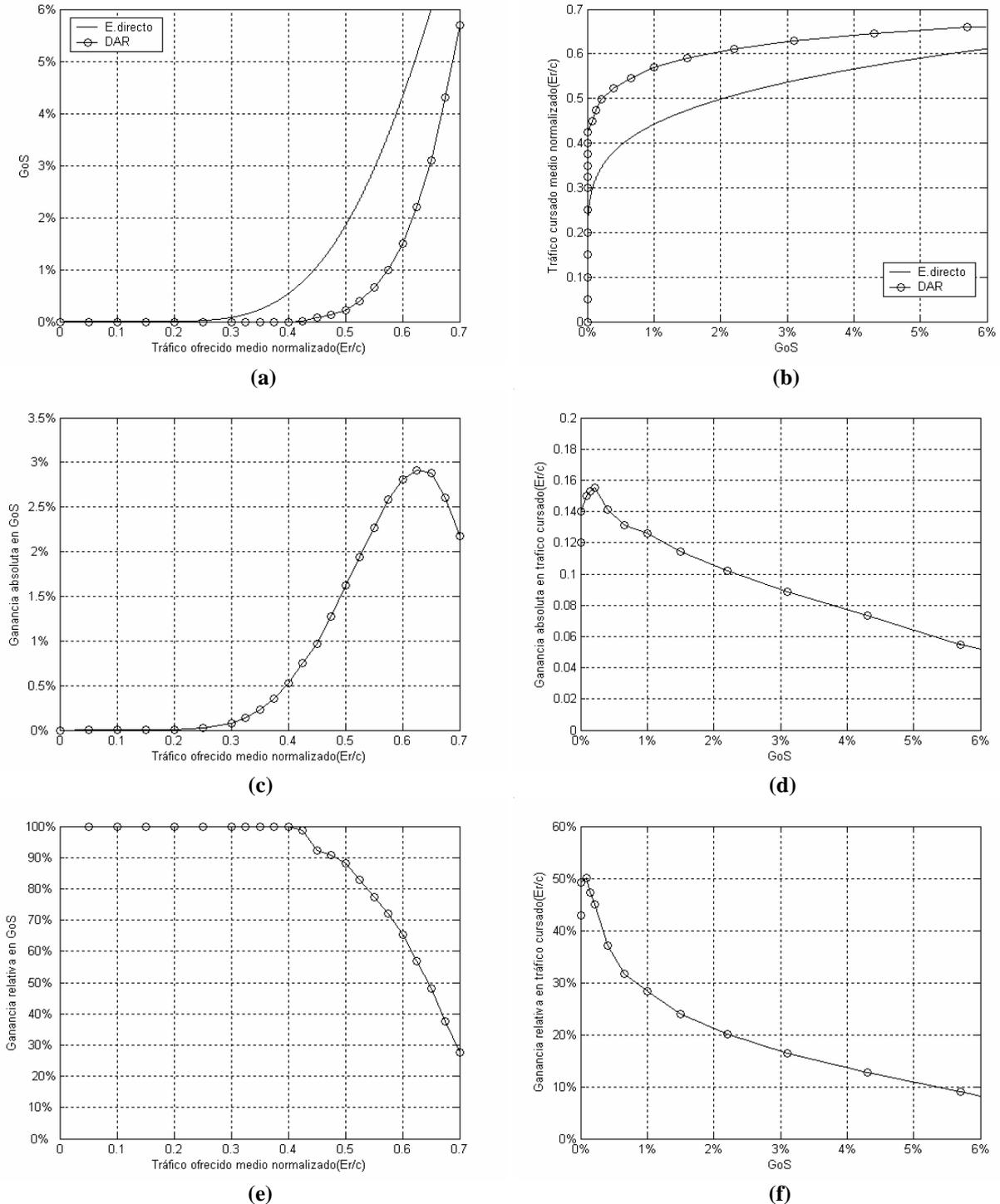


Figura 3.4 – Prestaciones de DAR en relación al encaminamiento directo. **(a)** GoS de la red para diferentes valores medios de tráfico ofrecido normalizado entre cada par de nodos. **(b)** Tráfico cursado para diferentes GoSs. **(c)** Ganancia absoluta en GoS. **(d)** Ganancia absoluta en tráfico cursado. **(e)** Ganancia relativa en GoS. **(f)** Ganancia relativa en tráfico cursado.

No obstante, no se va a trabajar en ninguno de los puntos de ganancia máxima mostrados en la tabla 3.1, sino en aquél en que el tráfico cursado sea mayor dentro de la cota máxima de GoS, un 1% en este proyecto. Es importante identificar este valor, así como calcular las ganancias que se obtienen, valores estos últimos que permitirán comparar las prestaciones de los diferentes algoritmos. Todos estos datos han sido calculados y son los que aparecen agrupados en la tabla 3.2.

G absoluta		
Prest.	Máximo en	G
GoS	A = 0,6260 Er/c	2,91%
Ac	GoS = 0,2177%	0,1536 Er/c

G relativa		
Prest.	Máximo en	G
GoS	A → 0 Er/c	100,00%
Ac	GoS = 0,0918%	50,14%

Tabla 3.1 – Ganancias absoluta y relativa máximas para GoS y para tráfico.

GoS (%)	Ac dir. (Er/c)	Ac DAR (Er/c)	G abs. (Er/c)	G rel. (%)
1	0,4417	0,5752	0,1335	30,2241

Tabla 3.2 – Tráficos y ganancias absoluta y relativa para GoS máximo.

3.4.2. Influencia del tamaño de la red

En el apartado anterior se han estudiado las prestaciones del DAR en GoS y en tráfico cursado para unos valores fijos de ciertos parámetros, que, como se verá a continuación, también influyen significativamente en las prestaciones del algoritmo. Así, se van a realizar simulaciones para diferentes parámetros relativos a la topología de la red, esto es, el número de nodos y la capacidad de los enlaces entre ellos. Recuérdese que siempre se considera una red simétrica totalmente conectada.

Dado que pueden presentarse inestabilidades al aumentar la capacidad de los enlaces, se emplea el método de reserva de troncales ya explicado para evitarlas, reservando para tráfico directo los circuitos que se indica en la tabla 3.3.

Por otro lado, estudiar los efectos de los dos parámetros mencionados (número de nodos y capacidad de los enlaces), requiere mantener constante el parámetro de simulación del apartado anterior, es decir, el tráfico ofrecido a cada enlace. Esto no puede hacerse directamente ya que uno de los parámetros de simulación es la capacidad de los enlaces, por lo que se ha optado por generar un tráfico ofrecido en cada caso tal que la misma red con encaminamiento directo presenta un GoS del 2%, concretamente los valores mostrados en la tabla 3.3. De esta forma se aumenta lógicamente el tráfico ofrecido al aumentar la capacidad de los enlaces, hecho que puede comprobarse en dicha tabla, así como un aumento implícito del tráfico ofrecido normalizado.

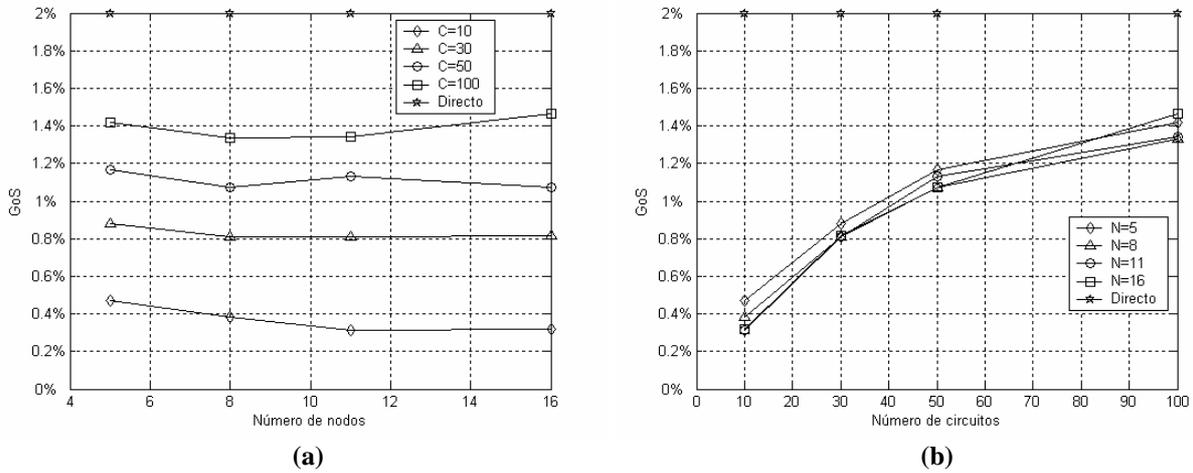
C	10	30	50	100
r	1	2	3	5
Ao (Er)	5,0840	21,9316	40,2551	87,9720
Ao (Er/c)	0,5084	0,7311	0,8051	0,8797

Tabla 3.3 – Número de circuitos reservados para tráfico directo, y tráfico ofrecido para alcanzar GoS=2% con encaminamiento directo.

A partir de las figuras 3.5 y 3.6 pueden extraerse varias conclusiones:

- En los márgenes estudiados de número de nodos y número de circuitos por enlace, DAR siempre mejora las prestaciones del encaminamiento directo.

- Es importante comprobar que DAR es poco sensible, como se observa en las figuras 3.5a y 3.6, al tamaño de la red en cuanto a número de nodos se refiere, lo que es una nueva ventaja que presenta.
- Las prestaciones del DAR son mejores para enlaces de menor velocidad, acercándose a las del encaminamiento directo para enlaces de mayor capacidad.



(a) (b)
Figura 3.5 – GoS del DAR para diferentes tamaños de red.

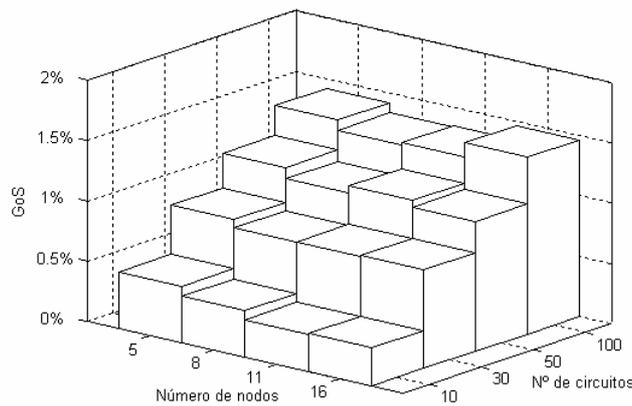


Figura 3.6 – GoS del DAR para diferentes tamaños de red.

3.5. Otros algoritmos aleatorios

Para finalizar el capítulo, se citarán varios algoritmos que también se basan en la idea de encaminamiento aleatorio. Recuérdese que se ha comenzado este capítulo justificando esta filosofía por ser una primera aproximación hacia el objetivo final de uniformización del reparto del tráfico en todos los enlaces de la red.

Cuando en cualquier disciplina se desea hacer un reparto uniforme en grupos puede emplearse la generación de variables deterministas uniformemente distribuidas, como el caso estudiado, o puede usarse un método de implementación aún más sencillo: el reparto cíclico de las unidades de los grupos. Particularizando al caso del encaminamiento, la elección de la ruta alternativa, tanto para intentar cursar cada petición cuando la directa está bloqueada como para la elección de una ruta alternativa válida hasta que ésta se bloquee, puede hacerse de forma cíclica, de forma que la última elegida pasa a ocupar la última posición de la lista de futuras elecciones. Un caso particular del encaminamiento cíclico es el encaminamiento CGH (llamado así por haber sido desarrollado por Chung, Graham y Hwang) que agrupa

las posibles rutas en grupos sobre los que se aplica el encaminamiento cíclico descrito, lográndose una mejora de las prestaciones.

Otro ejemplo es el STR (*State and Time-Dependent Routing*) que, básicamente, opera como el DAR de forma que las llamadas son primero ofrecidas a la troncal directa y, si ésta se bloquea, a la ruta alternativa; si también ésta se bloquea, la llamada se pierde, y se cambia la ruta alternativa activa. Este método es mixto en su distribución ya que trabaja de forma distribuida en los nodos de la red, pero recibe periódicamente un patrón de tráfico calculado por un nodo central. Este nodo también se encarga de restringir el conjunto de posibles rutas alternativas para cada par de nodos de conmutación de forma que $M < N - 2$. STR también emplea reserva de troncales para evitar inestabilidades en la red e integra técnicas de aprendizaje o *learning automata*. Es el sistema que empleó la NTT japonesa hasta el año 2002.

Como ya se comentó al comienzo del capítulo, el DAR surge como evolución del AAR (*Automatic Alternative Routing*). Este algoritmo se utilizaba en redes jerárquicas y, como ya se comentó, estaba basado en la búsqueda de la uniformización mediante la selección cíclica de una ruta entre todas las posibles. Las rutas directas estaban provistas de un número reducido de circuitos y cuando estaban completos, el tráfico desbordaba hacia uno de los posibles caminos que atravesaba un nodo *tándem* de jerarquía más alta. Por lo tanto, el tráfico en la ruta alternativa era la suma de los tráficos desbordados en todas las rutas directas. Sin embargo, se demuestra que si el tráfico que se ofrece a las rutas directas es aleatorio sólo se necesita una pequeña cantidad de circuitos en las rutas alternativas, menor al número de circuitos que habría que añadir a las rutas directas para reducir la probabilidad de bloqueo. Con la utilización de esta ruta alternativa también pueden aparecer problemas de inestabilidad que se pueden solucionar con la utilización de reserva de troncales.

Capítulo 4

Encaminamiento por ruta menos cargada.

LLR

4.1. Introducción

En el capítulo anterior se ha descrito un algoritmo de encaminamiento dinámico que busca la uniformización del tráfico en la red por su reparto mediante la elección aleatoria de las rutas alternativas. En su momento fueron explicadas las ventajas de este algoritmo de cara a su implementación. Básicamente, éstas se refieren a la sencillez para una implementación distribuida, ya que el conocimiento del estado de la red, concretamente de la ruta elegida como secundaria para cada par de nodos, no es explícito, sino que se deduce su inconveniencia a partir del bloqueo de una llamada que se le ofrece.

Parece lógico, pues, que un algoritmo de encaminamiento que haga uso de un conocimiento más profundo del estado de los enlaces pueda aprovecharlo para obtener mejores prestaciones. Así, la filosofía de encaminamiento descrita en este capítulo emplea la ruta con una carga cursada menor en el momento en que se produce la petición que se desea acomodar. Más exactamente, se encamina la llamada por la ruta con una menor función de coste, siendo ésta dependiente de forma proporcional a la carga que cursa.

4.2. Descripción del LLR (Least Loaded Routing)

4.2.1. Descripción general

El LLR o LBA (*Least Busy Alternative routing*) es un ejemplo de la filosofía antes descrita. En primer lugar, se intenta cursar la llamada por un circuito del enlace directo entre los nodos origen y destino, debido al menor uso de recursos. Si la troncal directa se encuentra bloqueada, la llamada es ofrecida a la ruta alternativa menos cargada, considerando una función de coste lineal dependiente del número de circuitos libres de cada ruta. Nótese que se busca la mayor capacidad libre y no la menor capacidad ocupada, hecho relevante en redes no simétricas. Dado que cada ruta secundaria está formada, como mínimo, por dos enlaces, se define la capacidad libre en una ruta como el mínimo de las capacidades libres en los dos enlaces que conforman la ruta.

En resumen, el LLR intenta cursar primero por la ruta directa y, en caso de estar ésta bloqueada, lo intenta por la ruta alternativa con máximo mínimo de las capacidades libres de los dos enlaces de la ruta.

Versiones simplificadas del LLR son las empleadas por algunas de las más importantes operadoras internacionales de comunicaciones, como el RTNR (*Real-Time Network Routing*) de la estadounidense AT&T, o el DCR (*Dinamically Controlled Routing*) de las canadienses Bell Canada y Stentor Canada.

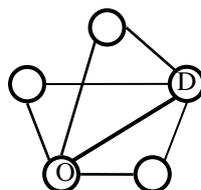


Figura 4.1 – LLR: todos los nodos son posibles nodos intermedios

4.2.2. Implementabilidad del LLR

Una de las versiones del LLR (concretamente el método DCR), que emplea alguna operadora, se implementa de forma centralizada, donde los nodos de conmutación informan periódicamente, usando la red de señalización, a un nodo central del estado de los enlaces a los que los nodos están conectados. Y, a su vez, el nodo central comunica ciertos patrones de tráfico que los nodos han de obedecer. Sin embargo, el LLR puro, que ha sido descrito anteriormente y que será estudiado en los apartados siguientes, puede ser implementado de forma distribuida. Téngase en cuenta que la información que un nodo necesita para decidir la ruta por la que cursar la llamada es el estado de dos grupos de enlaces:

- Aquellos a los que el propio nodo está conectado.
- Aquellos conectados a los nodos contiguos al nodo considerado.

En definitiva, es necesario conocer la capacidad libre de aquellos enlaces que pueden ser empleados como uno de los saltos de la ruta alternativa. Nótese que, en el caso de una red mallada, todos los enlaces entran en alguno de los dos grupos anteriores. El estado del primer grupo de enlaces es siempre conocido por cada nodo, mientras que para conocer cómo están los enlaces del segundo grupo es necesario que cada nodo comunique periódicamente la capacidad libre de sus enlaces entrantes y salientes a sus nodos vecinos, que son los que pueden utilizarlo como nodo intermedio, y de esta forma la matriz de estado de red quede actualizada.

La posible mejora de prestaciones, que se apoya en este más profundo conocimiento de la red, necesita como contraprestación la existencia de una red paralela de señalización o una capacidad residual, bien circuitos dedicados o bien paquetes de señalización, dentro de la propia red de comunicación, encareciendo, en ambos casos, el coste de la misma. Con el fin de analizar de forma práctica las prestaciones del LLR, se ha realizado la programación de un simulador en lenguaje C que implementa dicho algoritmo de encaminamiento. A partir de aquí, todos los resultados mostrados han sido obtenidos a partir de dicho simulador.

4.3. Prestaciones en redes multiclase

4.3.1. Distribución de los tiempos de servicio y entre llamadas

A continuación, se estudiarán las variaciones que la estadística del tiempo de duración del servicio o del tiempo entre llamadas provoca sobre las prestaciones de la red cuando el LLR es el algoritmo de encaminamiento empleado.

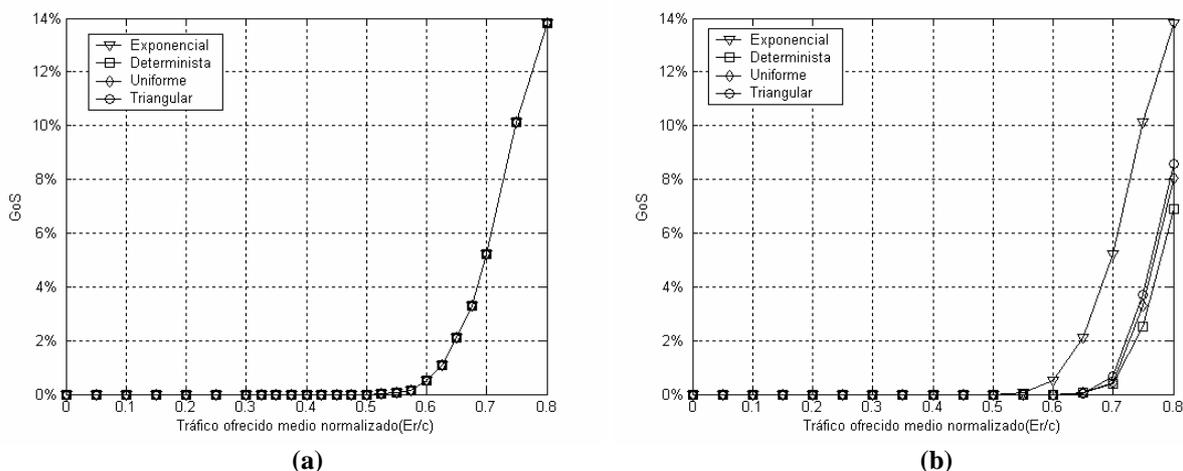


Figura 4.2 – Dependencia del GoS con la distribución del tráfico en una red totalmente conectada simétrica; $N=5$, $c=10$. (a) GoS para tiempo entre llamadas exponencial y tiempo de servicio variable. (b) GoS para tiempo de servicio exponencial y tiempo entre llamadas variable.

En la figura 4.2 se muestran de forma gráfica los resultados al modificar las distribuciones estadísticas de los tiempos mencionados, observándose de nuevo la baja sensibilidad del GoS ante variaciones en la estadística del tiempo de duración de la llamada, a diferencia de la del tiempo entre peticiones de conexión que afecta de la forma también vista anteriormente, empeorando las prestaciones para mayores valores de la varianza del tiempo. Los tiempos medios para los que se está estableciendo esta comparación son los mismo que en el DAR (figura 3.2).

4.4. Comparación de prestaciones con el encaminamiento directo

4.4.1. Resultados generales

A continuación se calcularán las ganancias en prestaciones del LLR relativas al encaminamiento directo y que ya se definieron en el capítulo anterior.

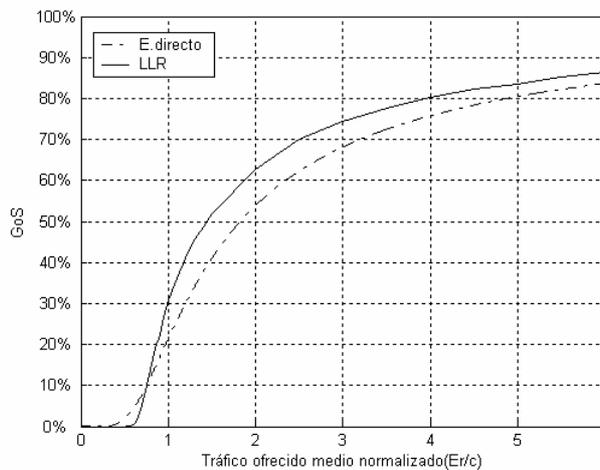


Figura 4.3 – Prestaciones del encaminamiento directo y del LLR en alta carga.

Antes de un estudio detallado de las ganancias que LLR logra en el margen de tráfico de trabajo, se han representado en la figura 4.3 las prestaciones del mismo junto con las del encaminamiento directo para un amplio margen de tráfico. Vuelve a comprobarse que la ganancia es positiva para tráficos menores de un determinado umbral, sobrepasado el cual, el sobreuso de rutas alternativas degrada las prestaciones, obteniéndose un GoS mayor que en el directo, por lo que la ganancia es negativa. Puede comprobarse que la separación entre las dos curvas es mayor que la que se tenía en el capítulo anterior para el DAR, ya que el LLR hace uso de todas las rutas alternativas posibles y el DAR sólo de una.

A continuación se estudiará de forma más completa el margen de tráfico en el que el GoS es suficientemente pequeño. En las figuras 4.4a a 4.4f se muestran gráficamente las comparaciones entre LLR y encaminamiento directo, en GoS, en tráfico cursado y en las ganancias absoluta y relativa respectivas, al igual que se hizo en el capítulo dedicado al DAR. En la figura 4.4a se comprueba que el LLR es capaz de cursar una relativa gran cantidad de tráfico sin presentar pérdidas. Podría fijarse un umbral aproximado de 0,5 Erlangs/circuito frente a los 0,3 Erlangs/circuito aproximados para el encaminamiento directo. Se explica así que la curva de ganancia absoluta, figura 4.4c, se corresponda con la del directo hasta los 0,5 Erlangs/circuito. Más allá de ese valor, LLR crece con una mayor pendiente por lo que su GoS llega a alcanzar al directo, cayendo, por lo tanto, la ganancia absoluta hasta hacerse negativa. El máximo se alcanza aproximadamente para los 0,63 Erlangs/circuito, donde se mejora el GoS en cuatro puntos porcentuales.

El hecho de que tanto el umbral mencionado en el párrafo anterior como que la pendiente de la curva de GoS sean mayores en el caso de LLR, hace que la máxima diferencia horizontal en la figura 4.4a

o vertical en la 4.4b se produzca cerca del origen, como puede observarse en la curva de ganancia de tráfico, en la figura 4.4d. El hecho de que los máximos absoluto y relativo se alcancen en el origen o en sus proximidades, es algo que también se verifica para el DAR, aunque no exactamente en el origen, ya que la curva de GoS del DAR parte con una pendiente menor que la del encaminamiento directo, si bien después crece hasta superarla y producirse el cruce de ambas curvas.

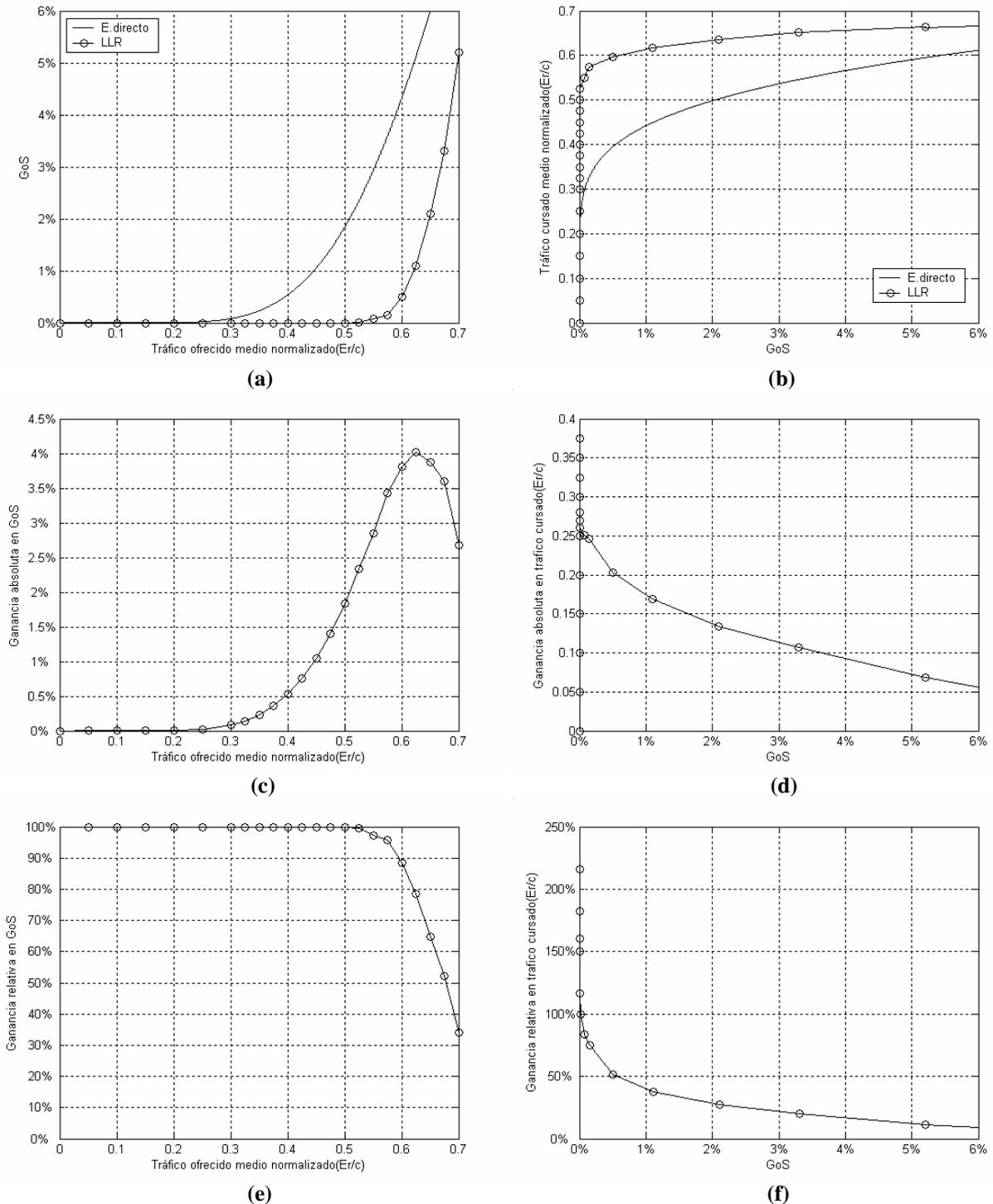


Figura 4.4 – Prestaciones de LLR en relación al encaminamiento directo. **(a)** GoS de la red para diferentes valores medios de tráfico ofrecido normalizado entre cada par de nodos. **(b)** Tráfico cursado para diferentes GoS. **(c)** Ganancia absoluta en GoS. **(d)** Ganancia absoluta en tráfico cursado. **(e)** Ganancia relativa en GoS. **(f)** Ganancia relativa en tráfico cursado.

Las ganancias máximas en GoS y en tráfico cursado, tanto absoluta como relativa, así como el punto en el que se producen, se presentan en la tabla 4.1. Por otra parte, las especificaciones del servicio obligan a trabajar por debajo de un GoS determinado, en este caso del 1%. En la tabla 4.2 se muestran los tráficos cursados por los encaminamientos directo y LLR, en las condiciones anteriores, así como la ganancia al emplear el segundo.

G absoluta		
Prest.	Máximo en	G
GoS	A = 0,6310 Er/c	4,0267%
Ac	GoS = 0%	0,375 Er/c

G relativa		
Prest.	Máximo en	G
GoS	A < 0,4 Er/c	100%
Ac	GoS = 0%	212,3808%

Tabla 4.1 – Ganancias absoluta y relativa máximas para GoS y para tráfico.

GoS (%)	Ac dir. (Er/c)	Ac LLR (Er/c)	G abs. (Er/c)	G rel. (%)
1	0,4417	0,6155	0,1736	39,3547

Tabla 4.2 – Tráficos y ganancias absoluta y relativa para GoS máximo.

4.4.2. Influencia del tamaño de la red

Para comprobar el efecto del tamaño de la red –en cuanto a número de nodos y capacidad de los enlaces que los unen– sobre las prestaciones del algoritmo, se han realizado las 16 mismas simulaciones que en el capítulo anterior para después realizar la comparación. Recuérdese que se trata de redes malladas simétricas, que se emplea el método de reserva de troncales con los valores de la tabla 3.3 y que los tráficos ofrecidos a cada red son tales que el GoS es del 2% para encaminamiento directo. Los resultados obtenidos son los que se muestran en las figuras 4.5 y 4.6.

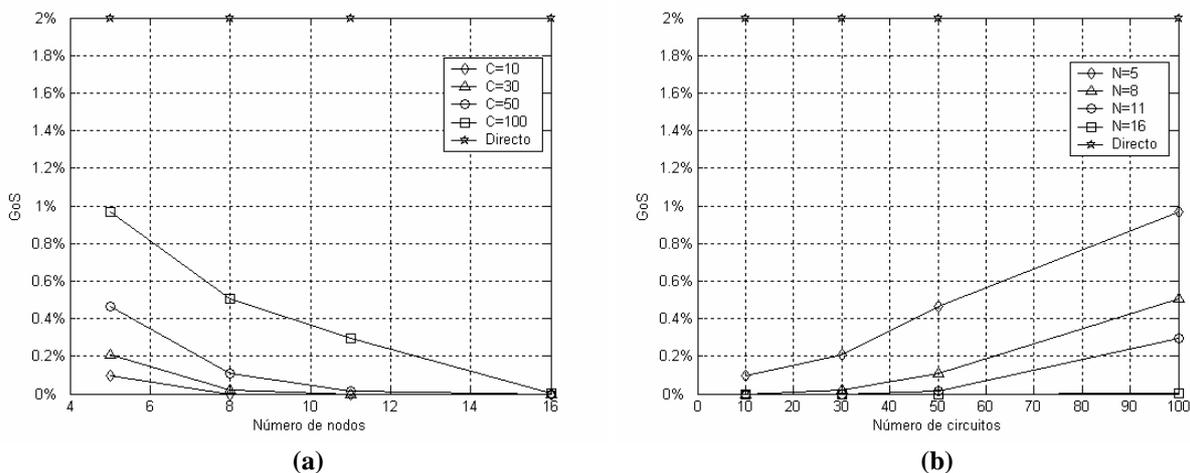


Figura 4.5 – GoS del LLR para diferentes tamaños de red.

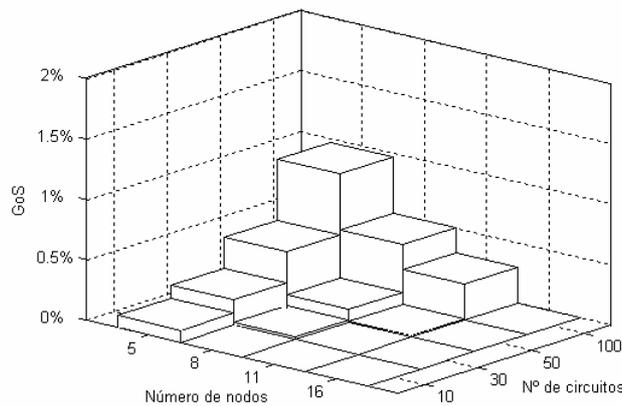


Figura 4.6 – GoS del LLR para diferentes tamaños de red.

A partir de las figuras 4.5 y 4.6 puede comprobarse que las prestaciones del LLR en los márgenes de N y c simulados son siempre mejores que para el encaminamiento directo, aunque la ganancia es dependiente de dichos parámetros. Así, se observa que las mejoras introducidas por el algoritmo son mayores para enlaces de pocas troncales y redes con un gran número de nodos. Lo primero puede deberse a diferentes causas, aunque uno de los motivos en los que habría que profundizar es en la reserva óptima de troncales. Lo segundo está motivado por el hecho de tener más rutas alternativas entre las que buscar la de menor coste.

4.5. Análisis del re-encaminamiento en LLR

Una de las mayores preocupaciones de los investigadores del encaminamiento dinámico es determinar la elección correcta del camino alternativo para las llamadas desbordadas. La idea es balancear la carga de tráfico entre los caminos alternativos tanto como sea posible, mientras que no caiga ninguno de ellos en congestión. A lo largo de los años, los investigadores han descubierto que el algoritmo LLR con reserva de troncales puede proporcionar una ganancia de *throughput* sobre el encaminamiento fijo, mientras que otros métodos más elaborados sólo pueden proporcionar una ganancia marginal.

En el encaminamiento dinámico, la decisión de encaminamiento se debe tomar en el momento de la llegada de la llamada, basándose en la información de red disponible en ese momento. Sin embargo, una vez tomada la decisión no se puede modificar. El re-encaminamiento es una práctica en la que las llamadas en las rutas alternativas pueden ser re-encaminadas al camino directo o a otra ruta alternativa menos congestionada. Este método trata de incrementar el *throughput* de los métodos tradicionales de encaminamiento dinámico, redistribuyendo la carga de la red para eliminar sobrecargas o embotellamientos.

El propósito del re-encaminamiento es redistribuir la carga de la red a lo largo del tiempo para liberar más capacidad para las llamadas a través de la ruta directa. Por ejemplo, si un canal ha sido liberado en el enlace AC debido a la finalización de una llamada, entonces una llamada que esté utilizando la ruta alternativa ABC puede ser reencaminada en ese momento al enlace AC. Haciendo esto conseguiremos liberar un canal en el enlace AB y otro en el BC y el *throughput* aumentará.

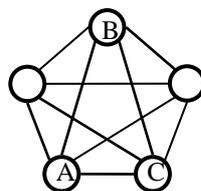


Figura 4.7 – Camino alternativo entre un par de nodos.

Aunque existen muchas versiones de re-encaminamiento, nosotros nos vamos a centrar en una versión implementada para el algoritmo LLR. Cuando se produce la llegada de una llamada, ésta será encaminada, en primer lugar, a su ruta directa X si hay un canal libre en esa ruta. Si la ruta X está bloqueada, una llamada alternativa que esté utilizando en ese momento la ruta X será elegida aleatoriamente y será reencaminada a su ruta directa, si es posible, para hacer hueco para la nueva llamada. Si el re-encaminamiento de esta llamada alternativa no tiene éxito, se elegirá otra llamada alternativa de la ruta X, y así sucesivamente. Si ninguna de las llamadas alternativas de la ruta X puede ser re-encaminada, el desbordamiento de la llamada que ha llegado se realizará hacia un camino alternativo con el máximo número de circuitos libres, es decir la ruta LLR. Si todos los caminos alternativos están completos, la llamada se bloqueará.

Los resultados numéricos de esta técnica muestran que el re-encaminamiento puede proporcionar un incremento significativo del *throughput* sobre el método LLR tradicional, bajo todas las condiciones. Un aspecto más interesante, es que el re-encaminamiento se muestra como un medio efectivo para mantener la estabilidad de la red con los métodos de encaminamiento dinámico. Por lo tanto, usando esta metodología, probablemente la reserva de troncales ya no sea necesaria y además se logra una ganancia de *throughput* adicional.

4.6. Otros algoritmos de ruta menos cargada

Se han mencionado al comienzo del capítulo algunos algoritmos de encaminamiento basados en la filosofía de cursar por la ruta menos cargada y que son empleados en la realidad por compañías de telecomunicaciones. Como se comentó en la introducción, estos algoritmos tienen en cuenta conceptos pertenecientes a varios niveles lógicos, por lo que son más complejos que los modelos que aquí se proponen, aunque sí consideran todos los conceptos que afectan al nivel de red.

Por otra parte, y debido básicamente a motivos tecnológicos, los algoritmos reales mencionados son más sencillos en el sentido de que no consideran todos los estados de ocupación de los enlaces, sino que definen varios *superestados* que agrupan a un cierto número de estados, de ahí que se denominen algoritmos ALBA (*Agregated Least Busy Alternative*). Las investigaciones realizadas demuestran que con un número reducido de *superestados* –un valor típico puede ser cuatro– las prestaciones que se logran son prácticamente las mismas que considerando todos los estados, pero reduciendo considerablemente el coste computacional del algoritmo, hasta el punto de hacerlo implementable con la tecnología actual.

4.6.1. Aplicación de las redes neuronales al encaminamiento dinámico en redes de conmutación de circuitos

A la hora de tomar las decisiones de encaminamiento se tienen que balancear las pérdidas y las ganancias, de tal forma que el rendimiento de los distintos componentes de la red se pueda optimizar. Esto se puede realizar con un controlador centralizado que recoja información del estado de la red y obtenga una decisión óptima global. Sin embargo, esta aproximación puede significar una carga computacional elevada para el controlador, y además es susceptible de provocar malos funcionamientos en la red.

Una aproximación alternativa es disponer de controladores descentralizados que tomen decisiones de encaminamientos óptimas localmente. Esto tiene las ventajas de menor carga computacional y robustez frente a las caídas de la red. Sin embargo, las decisiones de encaminamiento locales puede que no sean óptimas globalmente.

Las redes neuronales son candidatas para combinar las ventajas de ambas aproximaciones. Dado un conjunto de entrenamiento, con decisiones de encaminamiento óptimas globales, un controlador neuronal local puede aprender a inferir decisiones basándose en la información disponible localmente.

Así, el procedimiento consiste en dos partes. En la primera parte, se debe simular una red de conmutación de circuitos con un algoritmo de encaminamiento centralizado que consiga un rendimiento satisfactorio. Esto se puede utilizar para realizar el aprendizaje supervisado en los controladores neuronales locales. Para ello, se genera un conjunto de entrenamiento formado por numerosos ejemplos. Cada ejemplo se basa en la información de estado de la red y las decisiones de encaminamiento tomadas por el controlador centralizado. En la segunda parte, se construye una red neuronal para cada nodo. Sólo se dispone de la información local sobre el estado de la red en cada uno de los controladores neuronales. Después de realizarse el aprendizaje supervisado con el conjunto de entrenamiento, aprenderá a tomar decisiones basándose en dicha información local.

Esta técnica se ha empleado con la versión del algoritmo LLR utilizada en el método de encaminamiento dinámico RTNR, que implementó AT&T. También se ha empleado con este mismo algoritmo, pero utilizando el algoritmo *Simplex* de programación lineal para deducir las decisiones óptimas de encaminamiento.

Algunos ejemplos de la información con la que se construye el conjunto de entrenamiento para un nodo particular, son: el estado de los enlaces que parten de él, los intentos de llamada desde ese nodo, y los intentos de llamada totales. Hay que darse cuenta de que la única información global que necesita el controlador neuronal son los intentos de llamadas totales. En las redes reales, esta información puede ser difundida por la red de señalización.

Para cada nodo se utiliza una red neuronal de dos capas, cuyas entradas deben estar normalizadas para lograr la convergencia del conjunto de entrenamiento. A su vez, las salidas representan las decisiones de encaminamiento para cada una de las rutas. La primera capa ejecuta las tareas de clasificar los datos de entrada, y la segunda capa implementa las acciones de control de acuerdo con los pesos determinados por el clasificador. El aprendizaje se puede realizar tomando como función de error la suma de los errores cuadrados, y como algoritmo de optimización el descenso por gradiente.

En resumen, se puede decir que a partir de un algoritmo de control centralizado, las redes neuronales pueden aprender las decisiones que toma dicho algoritmo para cada una de las situaciones posibles. Ya que los intentos de llamada y el estado de los enlaces tienen una gran componente aleatoria, para realizar un buen aprendizaje es necesario generar un número de ejemplos suficientemente grande. Además, las redes neuronales tienen la ventaja de poder tomar decisiones rápidamente, y adaptarse a variaciones en la distribución del tráfico de entrada, de tal forma que se pueden generalizar para realizar el control en situaciones más extremas.

Capítulo 5

Comparativa entre los distintos métodos de encaminamiento dinámico

5.1. Introducción

Una buena estrategia de encaminamiento no sólo necesita poder disminuir el GoS de la red. Con el objetivo de poder trabajar bien en redes reales, se necesita que sea *simple*, para que miles de llamadas por segundo puedan ser encaminadas instantáneamente. También se necesita que sea *descentralizada*; un controlador central haría que el funcionamiento fuera extremadamente lento y que la red fuera desastrosamente mal si el controlador central sufriera un corte del suministro eléctrico o fuera desconectado del resto de la red.

Estos objetivos tan diferenciados se pueden conseguir simultáneamente. Hay esquemas descentralizados, y bastante simples, que pueden disminuir el GoS de la red, de forma tan eficiente como otros métodos más complejos. De esta forma, el DAR, aunque utiliza un modo de funcionamiento extremadamente simple, consigue un rendimiento bastante aceptable comparado con el que cualquier otro método de encaminamiento podría conseguir.

5.2. Comparación entre los tipos de encaminamiento

Como ya se comentó en el capítulo 2, los métodos de encaminamiento alternativo dinámico se pueden clasificar en tres grandes grupos: TDR, SDR y EDR, según aparece reflejado en las recomendaciones ITU-T. En este apartado vamos a resumir brevemente las características de cada uno de ellos y vamos a realizar un análisis de las bondades y desventajas de cada uno.

TDR

- Las tablas de encaminamiento están preestablecidas y se calculan de forma centralizada. En un instante de tiempo fijo del día o de la semana se modifican y se mantienen hasta la siguiente modificación.
- Una vez que se han cargado las tablas de encaminamiento los conmutadores no necesitan información adicional.
- Las reglas de selección de ruta están basadas en un encaminamiento secuencial.

SDR

- Las tablas de encaminamiento se alteran automáticamente de acuerdo al estado de la red, y se usa un periodo de actualización relativamente corto. Las tablas se pueden calcular de forma distribuida o centralizada, y periódicamente o llamada a llamada.
- Los conmutadores necesitan información sobre el estado de la red de forma periódica o bajo demanda.
- Las reglas de selección de ruta se modifican en respuesta al cambio del estado de la red y utilizan el principio de encaminar las llamadas por la mejor ruta disponible basándose en la información disponible del estado de la red.

EDR

- Las tablas de encaminamiento se actualizan de forma distribuida en caso de que la llamada falle o tenga éxito en una ruta determinada.
- Las tablas de encaminamiento se diseñan usando información de la red, obtenida durante el establecimiento de la llamada. Posteriormente no se necesita más información.
- La elección de la ruta alternativa se puede realizar aleatoriamente, cíclicamente, o de otras formas, y la ruta seleccionada se puede mantener mientras que las llamadas se establezcan con éxito en esa ruta.

Métodos centralizados frente a distribuidos

Las funciones básicas de encaminamiento se pueden realizar tanto de forma centralizada como de forma distribuida. El grado de descentralización depende del dinamismo, robustez y manejabilidad deseados para el sistema de encaminamiento.

En una implementación centralizada, una sola entidad es la encargada de realizar los cálculos de encaminamiento. Los procedimientos centralizados son fáciles de gestionar porque la funcionalidad reside en una sola entidad y, además, se permite una reducción de costes al concentrar el motor de cálculo en una sola localización.

Sin embargo, las implementaciones centralizadas tienen una serie de desventajas. Cuando una entidad central falla o es aislada del resto de la red, el rendimiento de la red entera se ve afectado. Este hecho se demostró en 1990 cuando un fallo en el procesador de encaminamiento de AT&T, provocó el rechazo de muchas llamadas durante un periodo de 9 horas. Por otra parte, la velocidad de respuesta de las funciones de encaminamiento a los cambios de estado de la red depende del estado de la central de cómputo. Una red muy cargada tendrá respuestas más lentas a los cambios de estado de la red que una red menos cargada. Este retraso en las respuestas de la red se produce justo cuando la red está más cargada y, por lo tanto, se necesitan respuestas más rápidas. Además, las porciones de la red que estén físicamente más distantes del procesador de encaminamiento también verán respuestas más lentas a los cambios de estado de la red. Por lo tanto, podemos decir que la concentración en las funciones de encaminamiento en un solo punto limita la velocidad de las respuestas de encaminamiento.

En un sistema de encaminamiento distribuido, múltiples entidades ejecutan las funciones de encaminamiento. Si la funcionalidad se replica, cada entidad proporciona funcionalidad de encaminamiento independientemente, y además, al ser distribuida, los conmutadores comparten información de estado para proporcionar dicha funcionalidad de forma compartida.

Aunque la descentralización requiere mayores costes y complicados esquemas de gestión de red, tiene numerosas ventajas. La replicación de la funcionalidad en múltiples entidades aumenta la fiabilidad del sistema. También proporciona respuestas a los cambios del estado local de la red, que un controlador centralizado no puede detectar. Dado que la carga computacional se extiende a lo largo de toda la red, la funcionalidad de encaminamiento se ve menos afectada por la sobrecarga de la red. Una implementación distribuida es también más escalable.

Para finalizar, podemos decir que a lo largo del tiempo la dependencia de una facilidad de cómputo centralizado ha causado errores masivos de encaminamiento bajo condiciones adversas de red, mientras que los algoritmos distribuidos se han presentado robustos frente a errores múltiples de nodos y enlaces en la red. Para redes de una sola clase de servicio los algoritmos de encaminamiento distribuidos han mostrado mejores rendimientos a un coste menor que los sistemas centralizados del mismo tamaño. Por ejemplo, los esquemas de encaminamiento distribuido como el DAR, se han manifestado superiores en coste y rendimiento a otros como el DNHR.

Por lo tanto, en este punto, entre todas las clases de encaminamiento dinámico, nos decantaríamos en primer lugar por los métodos del tipo EDR, a los que también deberíamos añadir aquellos métodos SDR que tienen una implementación distribuida.

Dependientes del tiempo frente adaptativos

Otra distinción importante en los sistemas de encaminamiento dinámico es la naturaleza de la información de estado utilizada para tomar las decisiones de encaminamiento. En las implementaciones dependientes del tiempo la elección de las rutas es un factor de la hora del día en que se realiza la petición. Esta aproximación proporciona predicciones del estado de la red en función del tiempo para evitar la congestión de enlaces o conmutadores en la red, ya que existe una fuerte correlación entre el tráfico de la red y la hora del día. Los métodos de encaminamiento dependientes del tiempo no se consideran adaptativos porque las alternativas de encaminamiento se mantienen fijas durante un periodo de tiempo constante (como por ejemplo una hora).

El encaminamiento adaptativo funciona guiándose por la información actual del estado de la red. El término adaptativo se refiere a la habilidad de la red para adaptarse a las distintas condiciones y encontrar una ruta óptima en las condiciones actuales. A diferencia del encaminamiento dependiente del tiempo, las rutas alternativas se evalúan en tiempo real y no suelen depender de la hora del día. El encaminamiento adaptativo requiere de un cómputo intensivo en los conmutadores, ya que la información de estado debe ser examinada frecuentemente, pero también proporciona mejores respuesta a las condiciones locales de la red.

Puesto que se está produciendo una transición en las redes de comunicaciones para proporcionar una gran variedad de servicios, la complejidad en las variaciones de tráfico también se está viendo incrementada. Los métodos de encaminamiento dinámico dependientes del tiempo, por sí solos, no pueden proporcionar la eficiencia necesaria para mantener los niveles de servicio. Por lo tanto, en este apartado volvemos a decidirnos por los métodos EDR y SDR, que realizan el diseño de las tablas de encaminamiento en tiempo real, frente a los métodos TDR.

Necesidad de información del estado de la red

Para poder conseguir un conocimiento exacto del estado de la red es necesario que los nodos se intercambien entre ellos mensajes periódicos de señalización a través de una red paralela de señalización o una capacidad residual dentro de la propia red de conmutación, con el incremento de costes que esto acarrea.

También hay que tener en cuenta que con el crecimiento de las redes, los flujos de información asociados al estado de la red también crecen de forma masiva. Por esta causa es aconsejable evitar los flujos masivos de información que presentan algunas formas de SDR, con el objetivo de hacer más escalables las redes. Por otra parte, los métodos EDR sólo requieren la información de red obtenida durante el establecimiento de la llamada, y posteriormente no necesitan más información. Por lo tanto, en este aspecto se prefieren los métodos EDR frente a los SDR.

Reglas de selección de ruta

La estrategia utilizada en los métodos EDR, y que consiste en que la ruta seleccionada se pueda mantener mientras que las llamadas se establezcan con éxito en esa ruta se conoce como *sticky-random*. Este método, como ya se dijo anteriormente, tiene un efecto de *empaquetamiento de llamadas* sobre la misma ruta alternativa, de forma que el tráfico desbordado de un par de nodos puede no interferir en el de otro par.

El principio anterior ha sido objeto de múltiples estudios durante los últimos 10 años. Las principales ventajas que tiene son su simplicidad de cómputo y los buenos rendimientos que se obtienen, comparables con los de otros algoritmos de encaminamiento dinámico más complejos.

Además, el hecho de que las nuevas rutas de encaminamiento se seleccionen en el momento en que una llamada se bloquea en la ruta alternativa actual se puede ver como un esquema de aprendizaje. En este sistema, la probabilidad de elegir un camino alternativo particular es de 1, si la llamada tiene éxito por la ruta alternativa actual, 0, si la llamada se ha bloqueado, y $1/(N-2)$ para el resto de posibles rutas alternativas. Para un largo periodo de tiempo todos los caminos alternativos habrán sido seleccionados un número igual de veces, ya que los nodos *tándem* son seleccionados aleatoriamente, y puesto que cada vez que una llamada falla se selecciona un nuevo nodo *tándem*, al final en cada ruta alternativa se habrá perdido un número igual de llamadas.

Por otro lado, en los métodos SDR el principal problema que se encuentra en las reglas de selección de ruta es la necesidad de una red de señalización; aunque pueden tomar decisiones de encaminamiento de una mayor precisión, puesto que disponen de información global de la red.

Por lo tanto, en este apartado se prefiere la estrategia utilizada en los métodos EDR por su gran simplicidad y el elevado rendimiento que se consigue.

5.3. Comparación DAR-LLR

En este apartado se pretende realizar una comparación minuciosa de las prestaciones obtenidas con los algoritmos de encaminamiento dinámico DAR y LLR. Hay que recordar que DAR es un método del tipo EDR y LLR es del tipo SDR. Se utilizará como base el estudio en profundidad realizado de estos dos algoritmos en los capítulos 3 y 4.

5.3.1. Comparación descriptiva

Para comenzar, hay que mencionar que las estrategias utilizadas por DAR y LLR para la selección de la ruta alternativa son completamente diferentes. DAR busca la uniformidad de la distribución del tráfico en los nodos y enlaces que componen la red mediante la generación de una variable aleatoria discreta uniforme, que se va a hacer corresponder con una ruta alternativa. Por su parte, LLR encamina la llamada por la ruta con una carga cursada menor en el momento en que se produce la petición.

Ambos algoritmos siempre intentan cursar en primer lugar la llamada por un circuito del enlace directo entre los nodos origen y destino, debido al menor uso de recursos. Sin embargo, en el DAR a diferencia del LLR, el encaminamiento por la ruta alternativa es fijo para cada par de nodos hasta que esta ruta se bloquea. En el LLR, como ya se ha comentado, la ruta alternativa es aquella que en cada momento posee una mayor capacidad residual.

En cuanto a la implementación, hay que decir que el DAR tiene un menor coste computacional que el LLR, hecho que ha dado lugar a que las implementaciones en redes reales posean versiones simplificadas del LLR, y no implementaciones totales. Una de estas simplificaciones, como vimos en el capítulo 4, se encuentra en la definición de *superestados*.

Otra característica importante es que el DAR desconoce el estado exacto de la red, mientras que el LLR hace uso de un conocimiento más profundo del estado de los enlaces. Una consecuencia importante del hecho anterior es que en el DAR no es necesario que los nodos se intercambien entre ellos mensajes periódicos de señalización, mientras que en el LLR es necesaria la existencia de una red paralela de señalización o una capacidad residual dentro de la propia red de comunicación.

También hay que destacar que en el DAR cada nodo de la red toma decisiones de encaminamiento independientemente del resto de nodos, evitando la necesidad de disponer de un nodo central que gestione todas las tareas de encaminamiento. En el LLR, por el contrario, aunque también puede ser implementado de forma distribuida, una de las implementaciones en una red real funciona de

forma centralizada, y en este caso también deben de existir mensajes de señalización desde el nodo central hacia los nodos de conmutación informando de la decisión tomada.

De cara a los costes, la alternativa por la que opta el DAR evita, por una parte, los costes del nodo central y, mucho más importante, los costes de una red de señalización asociada.

5.3.2. Comparación de prestaciones con el encaminamiento directo

En primer lugar, si observamos la figura 5.1, donde se reflejan las prestaciones de ambos métodos de encaminamiento frente al encaminamiento directo para un amplio margen de tráfico, se comprueba que en ambos casos la ganancia es positiva para tráficos menores de un determinado umbral, sobre el cual el sobreuso de rutas alternativas degrada las prestaciones, obteniéndose un GoS mayor que en el directo.

Sin embargo, si observamos las diferencias entre ambas curvas (ver figura 5.2) notamos, en primer lugar, que la separación de la curva del LLR con la del encaminamiento directo es mayor que en el caso del DAR, ya que el LLR hace uso de todas las rutas alternativas posibles y el DAR sólo de una. Esto hace que debajo del umbral, anteriormente mencionado, las prestaciones del LLR sean bastante mejores que las del DAR y por encima de dicho umbral ocurra lo contrario, el DAR sea mejor que el LLR. Esto último no tiene mucha importancia porque ocurre fuera del margen de tráfico en el que el GoS es suficientemente pequeño, que es el margen de interés. También se puede observar que el umbral anterior, donde se cruzan la curva del método de encaminamiento dinámico y la curva del encaminamiento directo, ocurre en ambos casos, DAR y LLR, aproximadamente para el mismo valor de tráfico ofrecido medio normalizado.

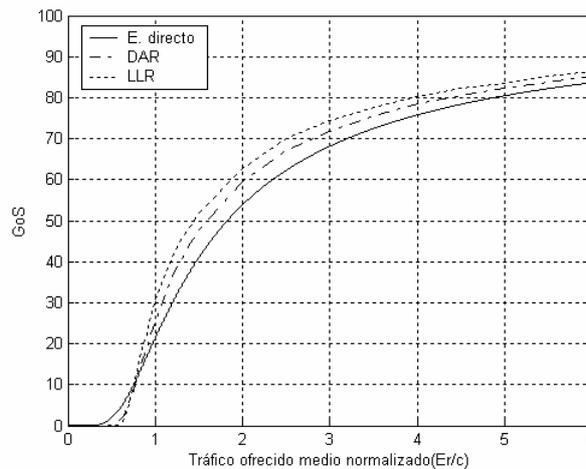


Figura 5.1 – Prestaciones del encaminamiento directo, del DAR y del LLR en alta carga.

A continuación, se va a realizar una comparación en profundidad entre las figuras 3.4 y 4.4. Si se observa la figura 5.2, que representa el GoS de la red, se comprueba que el LLR es capaz de cursar una relativa gran cantidad de tráfico sin presentar pérdidas. Se podría fijar un umbral aproximado de 0,5 Erlangs/circuito frente a los 0,4 Erlangs/circuito aproximados para el DAR. Sin embargo, dado que, como se dijo anteriormente, el punto en el que se cruzan las curvas del método de encaminamiento dinámico y del encaminamiento directo coincide aproximadamente para ambos casos, esto trae consigo que la curva del LLR deba crecer con una mayor pendiente que la del DAR.

Si comparamos las curvas de ganancia absoluta en GoS (ver figura 5.3) observamos una curiosidad y es que en ambos métodos esta curva presenta una forma similar, de hecho el máximo valor de ganancia se da en ambos casos para un tráfico ofrecido entre 0,62 y 0,63 Erlangs/circuito. Sin embargo, como era de esperar, el valor de la ganancia absoluta para el LLR (4,0267%) es mayor a la del DAR (2,91%). De la misma forma, las curvas de ganancia relativa en GoS (ver figura 5.4) para ambos métodos también presentan una forma bastante similar; de hecho, en ambos casos el máximo se produce

para un tráfico ofrecido de 0 Erlangs/circuito. Sin embargo, el margen de ganancia relativa máxima en GoS para el LLR tiene una mayor longitud, extendiéndose hasta los 0,5 Erlangs/circuito, mientras que en el DAR sólo llega, aproximadamente, hasta los 0,4 Erlangs/circuito.

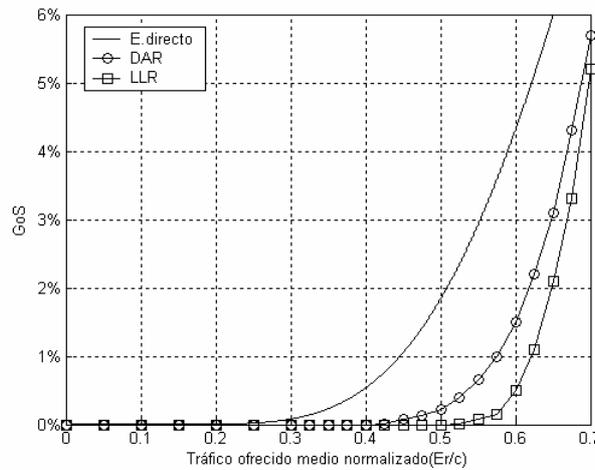


Figura 5.2 – GoS de la red para diferentes valores medios de tráfico ofrecido medio normalizado.

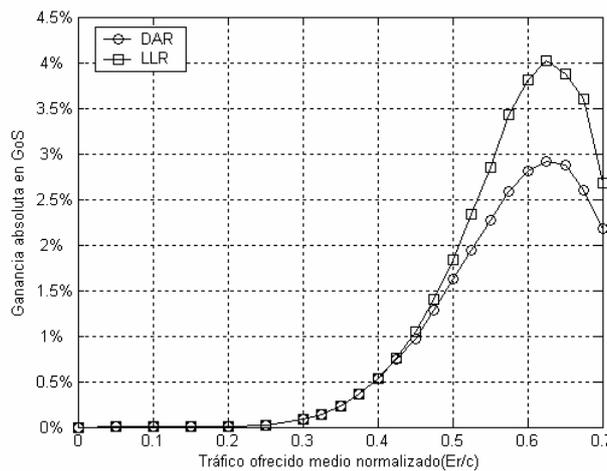


Figura 5.3 – Ganancia absoluta en GoS de la red para diferentes valores medios de tráfico ofrecido medio normalizado.

El hecho de que tanto el umbral a partir del cual aparecen pérdidas sea mayor en el LLR que en el DAR, como que la pendiente de crecimiento en el LLR sea mayor que en el DAR, hacen que la máxima diferencia vertical en la figura 5.5, que refleja los tráficos cursados para diferentes GoSs, para el LLR se produzca más cerca del origen y con un valor mayor que para el DAR. De hecho, si comparamos las curvas de ganancia absoluta en tráfico cursado, en la figura 5.6, observamos que las dos presentan una forma similar; sin embargo, el valor máximo para el LLR se produce en el origen (GoS = 0%), con un valor de 0,375 Erlangs/circuito, mientras que para el DAR se produce para un GoS = 0,2177%, con un valor de 0,1536 Erlangs/circuito. De la misma forma, si comparamos las curvas de ganancia relativa en tráfico cursado, en la figura 5.7, observamos un comportamiento similar al caso anterior, con una ganancia máxima del LLR, y superior a la del DAR, en el origen (GoS = 0%).

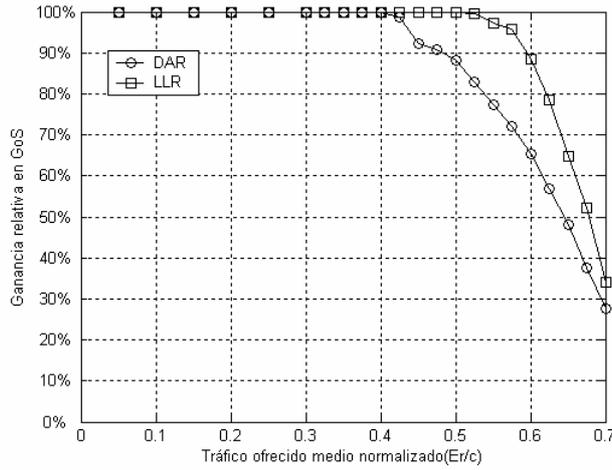


Figura 5.4 – Ganancia relativa en GoS de la red para diferentes valores medios de tráfico ofrecido medio normalizado.

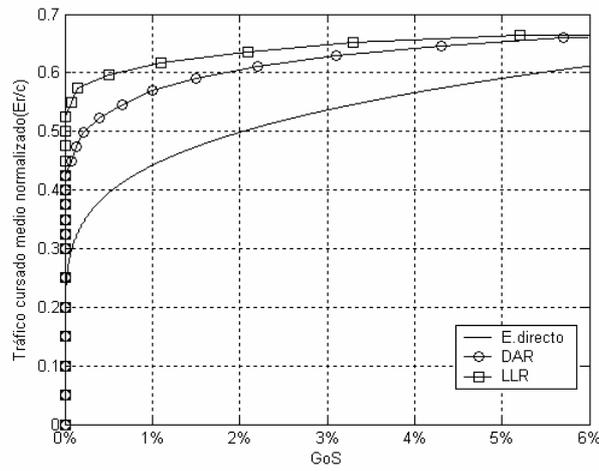


Figura 5.5 – Tráficos cursados para diferentes GoSs.

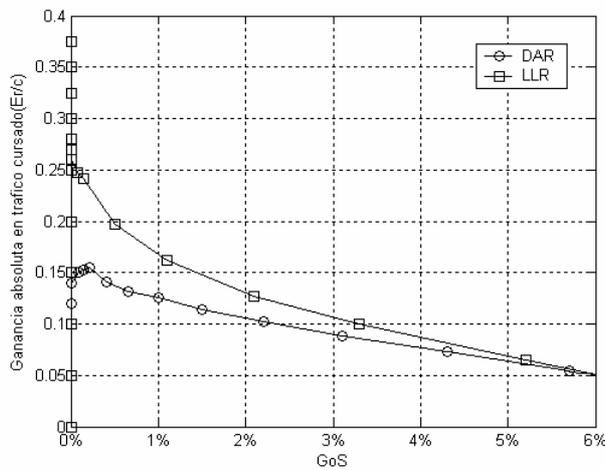


Figura 5.6 – Ganancia absoluta en tráfico cursado.

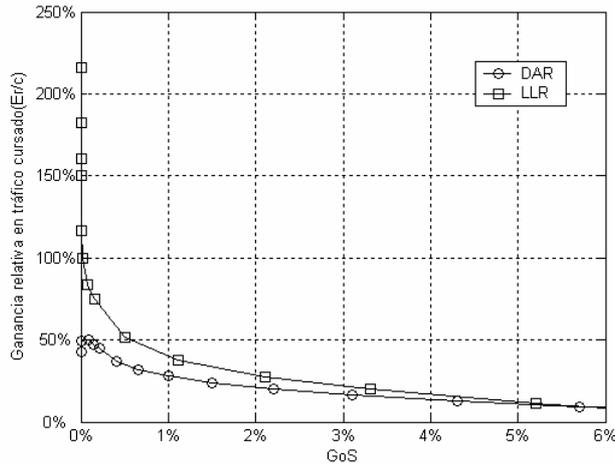


Figura 5.7 – Ganancia relativa en tráfico cursado.

Por último, si comparamos los valores que aparecen en las tablas 3.2 y 4.2, que reflejan los valores del tráfico cursado para un GoS máximo del 1%, observamos, como era de esperar, que el tráfico cursado para el LLR es mayor que para el DAR y, por lo tanto, también son mayores los valores de las ganancias absoluta y relativa.

5.3.3. Influencia del tamaño de la red

Si comparamos las figuras 5.8 y 5.9 observamos que en los márgenes estudiados de número de nodos y número de circuitos por enlace, DAR y LLR siempre mejoran las prestaciones del encaminamiento directo. Además, se presenta una analogía para ambos métodos en la influencia que el número de circuitos por enlace tiene sobre las prestaciones de la red, pues para ambos algoritmos las prestaciones obtenidas son mejores para enlaces de menor velocidad, acercándose a las del encaminamiento directo para enlaces de mayor capacidad. Aunque eso sí, las prestaciones obtenidas por el LLR son siempre superiores a las obtenidas con el DAR.

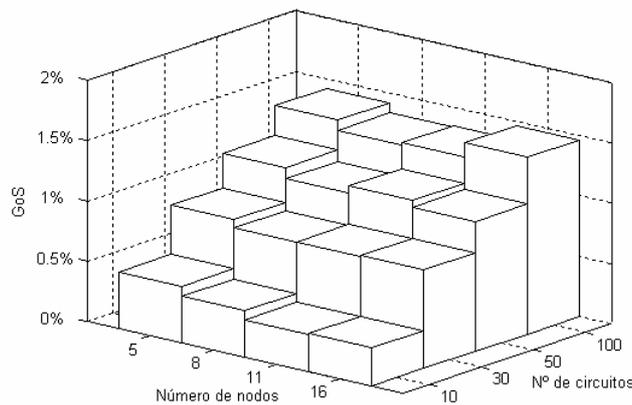


Figura 5.8 – GoS del DAR para diferentes tamaños de red.

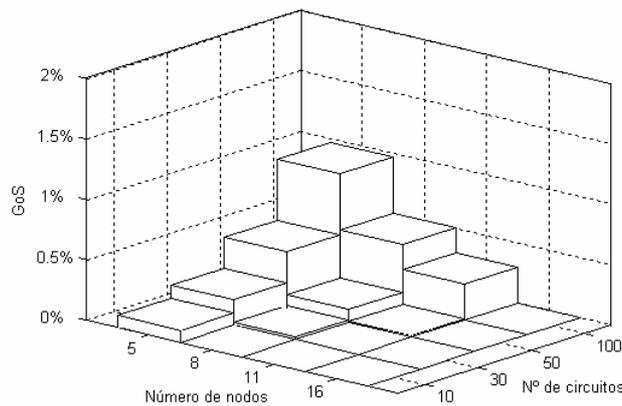


Figura 5.9 – GoS del LLR para diferentes tamaños de red.

Sin embargo, podemos observar que el número de nodos de la red influye de manera distinta en ambos algoritmos, pues mientras que el DAR es poco sensible a la variación de este parámetro, en el LLR se introducen mayores mejoras en redes con un gran número de nodos, ya que hay más rutas alternativas entre las que buscar la de menor coste.

De esta forma, las mejores prestaciones para el LLR se obtienen para redes con un número de nodos elevado y un número de circuitos por enlaces más reducido. Mientras que las peores prestaciones se obtienen para redes con un número de nodos reducido y un número de circuitos por enlace elevado.

En el DAR, las mejores prestaciones también se obtienen para redes con un número de circuitos por enlace reducido, y las peores en redes con un número de circuitos por enlace elevado; en ambos casos, con independencia del número de nodos de la red.

5.3.4. Distribución de los tiempos de servicio y entre llamadas

Interesa estudiar cuán robustos son el DAR y el LLR ante variaciones en las distribuciones del tiempo entre llegadas y del tiempo de duración de la conexión. Esto dará una idea de la robustez del DAR y del LLR ante la introducción de nuevos servicios de modelos de tráfico aún desconocidos.

Si comparamos los resultados reflejados en las figuras 3.2 y 4.2, para el DAR y el LLR, respectivamente, observamos en primer lugar que para ambos métodos la distribución estadística del tiempo de conexión apenas influye, mientras que sí lo hace la del tiempo entre peticiones. Apreciamos que ambos algoritmos muestran mayores valores de GoS para una distribución exponencial del tiempo entre llamadas, y menores para una distribución determinista, por la mayor varianza de la exponencial y varianza nula de la determinista. La razón por la cual la distribución del tiempo de conexión no influye en los valores obtenidos en el GoS ya se comentó en el punto 3.3.1, y viene provocada por el hecho de que en los sistemas sin cola el tiempo residual de servicio medio es independiente de la distribución estadística que adquiera el tiempo de conexión.

Sin embargo, las diferencias que presentan las gráficas para ambos métodos se deben a las razones ya expuesta anteriormente. Éstas son, en primer lugar que el LLR presenta un mayor umbral de tráfico ofrecido que el DAR para que aparezcan pérdidas. Por otra parte, la pendiente con la que crecen las pérdidas es mayor para el LLR que para el DAR.

5.4. Conclusiones

Podemos extraer varias conclusiones de la comparación entre los distintos métodos de encaminamiento llevada a cabo en este capítulo. En primer lugar, han quedado demostradas las ventajas de utilizar un método de encaminamiento distribuido frente a uno centralizado. Dentro de este tipo de métodos de encaminamiento encontramos los métodos EDR, y aquellos SDR distribuidos.

En segundo lugar, también se han demostrado las ventajas de los métodos adaptativos frente a aquellos que tienen un comportamiento dependiente del tiempo, por lo que también se prefieren los métodos EDR y SDR, frente a los TDR.

Se han presentado los inconvenientes que trae consigo el requerir información global sobre el estado de la red para ejercer las funciones propias de los métodos de encaminamiento dinámico. Estos inconvenientes se refieren sobre todo a los costes añadidos de una red de señalización, y a las dificultades de escalabilidad que tienen los métodos que necesitan información de los distintos nodos de la red para su funcionamiento, como son los métodos SDR.

Por último, se ha demostrado que las reglas de selección de ruta basadas en el principio de *sticky-random*, como las que utilizan algunos métodos EDR, además de su gran simplicidad y la ausencia de necesidad de una red señalización, logran unos rendimientos comparables con los de otros algoritmos de encaminamiento dinámico más complejos.

Por lo tanto, podemos decir que aunque el rendimiento obtenido en algunos métodos SDR es superior al conseguido con métodos EDR, la gran simplicidad de los últimos, y sobre todo el reducido coste implantación, hace que nos podamos decidir por un método de este tipo para implantarlo en una red real. Sin embargo, puede ser que para algunas aplicaciones el coste no sea un problema, sino que el principal objetivo sea el de conseguir un rendimiento elevado. En este caso, se recomienda la utilización de métodos del tipo SDR.

Con las simulaciones realizadas de los algoritmos DAR y LLR se comprueban las mejores prestaciones del LLR, que es un método SDR. Sin embargo no son muy superiores a las del DAR, que es un método EDR, y además dependen del margen de tráfico ofrecido y GoS en el que trabajamos, tal y como se observa en las curvas de prestaciones y de ganancias absolutas. Pueden haber situaciones en las que no nos importe conseguir un rendimiento ligeramente inferior, si a cambio consiguen reducir los costes y los problemas de escalabilidad asociados al LLR. Por ello, hay que decir que la utilización en redes reales del DAR, a pesar de su gran simplicidad, se puede justificar de una forma muy razonada, ya que logra conseguir unos rendimientos muy aceptables a unos costes muy inferiores.

Por último, se puede comprobar que algunos parámetros, como la velocidad de los enlaces de la red, tienen una influencia similar en el DAR y en el LLR, por lo que no influyen como un factor determinante a la hora de decantarnos por uno u otro método.

Capítulo 6

Aplicación del encaminamiento dinámico a redes de conmutación de paquetes

6.1. Introducción

Se ha aprendido mucho de las implementaciones de encaminamiento dinámico en redes de conmutación de circuitos, y las lecciones y la experiencia adquiridas en este campo han contribuido a la evolución de otras tecnologías. De hecho, algunos principios de encaminamiento dinámico fueron analizados e implementados en el marco de la primera generación de redes de datos como Arpanet y X.25. A continuación se muestra un resumen de las lecciones adquiridas:

- Uso de métodos EDR (*Event-Dependent Routing*) como alternativa a los métodos TDR (*Time-Dependent Routing*). Además, estos métodos evitan los flujos masivos de información de estado asociados con algunos métodos SDR, con el objetivo de hacer más escalables a las redes.
- Uso de reserva de ancho de banda dinámico, para hacer a las redes más estables y eficientes.
- Uso de los principios de clase de servicio, para habilitar la asignación de ancho de banda dinámico para clases de servicio individuales.

Algunos de estos principios se han extendido al entorno de las redes de conmutación de paquetes, pues se han considerado en redes ATM, y también son extensibles a la Ingeniería de Tráfico dentro de las redes IP basadas en MPLS. Las redes actuales están evolucionando rápidamente para poder transportar una gran variedad de servicios en redes basadas en el protocolo IP, conducido en parte por el crecimiento extremadamente rápido de los servicios de datos basados en este protocolo. Hay una serie de requisitos de rendimiento que normalmente deben alcanzar las distintas clases de servicio, como el bloqueo extremo a extremo, el retardo o el *jitter*. En este contexto, el encaminamiento juega un papel esencial para conseguir las restricciones de calidad asociadas a los diferentes servicios. Por ello, los métodos de encaminamiento dinámico se incluyen en el establecimiento de conexión para servicios multimedia de banda estrecha y de banda ancha dentro de redes multiservicio y entre redes multiservicio. Aquí, una red multiservicio se refiere a aquella en la que varias clases de servicio comparten la transmisión, la conmutación, la gestión y otros recursos de la red, como es el caso de la red Internet.

6.2. Redes ATM

El problema del encaminamiento de circuitos virtuales en redes ATM es un tema similar al encontrado en redes de conmutación de circuitos. Por ello, se han propuesto algunas extensiones de métodos de encaminamiento en redes de conmutación de circuitos para aplicarlos en este tipo de redes. El objetivo de estos esquemas de encaminamiento normalmente sólo considera un requisito de calidad de servicio (QoS), que es el ancho de banda efectivo que consiguen.

La primera aproximación fue la de extender el *encaminamiento alternativo con reserva de troncales* a una topología general. Bajo este esquema, se usa un protocolo de encaminamiento vector-distancia para construir las tablas de encaminamiento de cada nodo, con múltiples caminos disponibles para alcanzar cada posible nodo destino. El establecimiento de un circuito virtual primero se encamina hacia el camino primario (el directo). Si no hay disponibles suficientes recursos a lo largo de este camino, se intentan caminos alternativos en orden creciente de número de saltos hacia el destino. Un *flag* en el paquete de establecimiento de circuito virtual indica si el camino que se ha elegido es el directo o uno alternativo, y a continuación se reserva el ancho de banda en los enlaces de la ruta seleccionada, teniendo

en cuenta las restricciones de la reserva de troncales. Los valores de la reserva de troncales se determinan basándose en datos predeterminados sobre las características del tráfico.

La siguiente idea fue la de importar la noción del *encaminamiento alternativo controlado*. Para hacer esto, en primer lugar a cada circuito virtual se le asigna un coste de encaminamiento máximo permisible. Este coste se puede establecer basándose en el tráfico esperado a transportar o simplemente basándose en el camino más corto hacia el destino. A continuación, cada enlace se asocia con una métrica, que crece exponencialmente con la utilización de dicho enlace. Un conmutador calcula un camino viable de coste mínimo para un circuito virtual basándose en la métrica de los enlaces y en los requisitos de calidad de servicio del circuito virtual. El circuito virtual se admite si el coste del camino es menor o igual al máximo permisible.

Estas aproximaciones están incluidas en el protocolo PNNI (*Private Network-Network Interface*) del Forum ATM. Se presenta una lista parcial de sus principales características:

- Escalable a redes muy grandes.
- Soporte de encaminamiento jerárquico.
- Proporciona encaminamiento dinámico.
- Soporte de QoS.
- Soporta múltiples métricas y atributos.

La especificación PNNI se subdivide en dos protocolos: uno de señalización y otro de encaminamiento. El protocolo de señalización PNNI se usa para establecer conexiones punto a punto y punto a multipunto, y soporta encaminamiento alternativo. Se permite que cada implementación utilice su propio algoritmo para el cálculo del camino.

El protocolo de encaminamiento PNNI es un protocolo dinámico que proporciona información sobre la topología de la red. La información de la topología es relativa al conjunto de recursos (nodos, enlaces y direcciones) que definen la red. Los recursos se califican utilizando un conjunto de métricas y atributos (retardo, ancho de banda disponible, *jitter*, etc.), y se agrupan en clases de tráfico soportado.

6.3. Redes IP

En [5] se proporciona un análisis de las prestaciones de pérdida/retraso de tráfico, y control de carga, para varios métodos de encaminamiento dinámico implementados en redes IP. Estos métodos controlan la respuesta de una red a las demandas de tráfico y otros estímulos, como desbordamiento de tráfico, y fallos en nodos o enlaces. A partir de los resultados de estos estudios, se propusieron los métodos de encaminamiento y de control de acceso más importantes para tenerlos en cuenta en la evolución de las tecnologías IP.

En las redes IP, hay a menudo una relación jerárquica entre diferentes áreas o subredes. En estas topologías de multidominio jerárquico se usan normalmente protocolos de encaminamiento IP del tipo de OSPF (*Open Shortest Path First*) y BGP (*Border Gateway Protocol*). Este tipo de protocolos los podemos relacionar directamente con los métodos de encaminamiento alternativo fijo.

Por su parte, los distintos tipos de métodos de encaminamiento alternativo dinámico (TDR, SDR y EDR) los podemos ver utilizados de forma conjunta con métodos de reserva de ancho de banda en redes IP. También los podemos ver aplicados a la gestión de las tablas de encaminamiento. En el estudio anteriormente mencionado se encontró que los métodos EDR (*Event-Dependent Routing*) tienen un rendimiento tan bueno o mejor que los métodos SDR en la gestión de los caminos MPLS, a lo que hay que añadir que estos métodos de selección de camino pueden mejorar de forma significativa la escalabilidad de la red. En los métodos EDR no se necesita la información del ancho de banda disponible en cada uno de los enlaces para calcular los distintos caminos, por lo tanto se evita el flujo de este tipo de información, reduciendo la sobrecarga. Como ya se comentó en el capítulo 2, los métodos EDR son

adaptativos y distribuidos, y típicamente utilizan modelos de aprendizaje para aprender buenos caminos, como el método utilizado en DAR.

A continuación, se muestran algunas de las conclusiones –extraídas del trabajo anterior– sobre el rendimiento obtenido con la utilización de los métodos de encaminamiento dinámico en redes IP:

- Con la utilización de métodos de encaminamiento dinámico siempre se alcanza un rendimiento mejor en la red, respecto a la no utilización de dichos algoritmos.
- Los métodos de encaminamiento dinámico EDR muestran un rendimiento mejor o igual al de los métodos SDR.
- En topologías planas con una sola área proporcionan un mejor rendimiento, en comparación con topologías multiárea jerárquicas.
- Los distintos métodos de encaminamiento dinámico utilizados en una misma red pueden interactuar entre ellos.

6.4. MPLS

Uno de los grandes desafíos en las modernas redes de comunicaciones es transportar diferentes tipos de información entre los usuarios, asegurando una cierta calidad de servicio o QoS (*Quality of Service*). Estas redes deberán ser diseñadas para poder transportar una multiplicidad de servicios distintos con requisitos heterogéneos y soportar futuros servicios con requisitos todavía desconocidos. Este escenario diverge enormemente de la situación planteada hasta hace poco, donde para cada servicio existía una red dedicada. Estas circunstancias exigen nuevas maneras de plantear el control y gestión del tráfico cursado.

Para poder proporcionar calidad de servicio se precisa trabajar, a diferencia de la red Internet que trabaja en modo datagrama, con tráfico orientado a conexión, de forma similar a como han trabajado las redes telefónicas tradicionales. Sin embargo, en contraste con estas últimas, donde se transporta una única clase de tráfico, las nuevas redes deben ser capaces de transportar un tráfico heterogéneo, donde se superpone tráfico con requisitos estrictos de retardo y *jitter* (como es el caso de la voz o el vídeo), con tráfico de datos donde los parámetros de retardo son secundarios y el principal requisito a satisfacer es la entrega de los datos de forma correcta y sin errores.

Por lo tanto, para conseguir que los diversos algoritmos y protocolos de encaminamiento mejoren el rendimiento de Internet y que la ruta encontrada cumpla múltiples requisitos de calidad de servicio, se precisa un encaminamiento en origen y orientado a conexión. De hecho, los últimos trabajos presentados utilizan DiffServ (*Servicios Diferenciados*) sobre MPLS (*Multi-Protocol Label Switching*) para conseguir que los algoritmos de encaminamiento sean capaces de encontrar caminos con calidad de servicio y permitir un adecuado aprovechamiento de los recursos.

El encaminamiento orientado a conexión consigue distribuir el tráfico entre los distintos recursos de la red, ya que, de esta manera, se desvía únicamente el tráfico que causa la sobrecarga y no el que ya ha sido correctamente encaminado y para el cual hay suficientes recursos.

La mayoría de los protocolos de encaminamiento tradicionales en redes IP, como OSPF, no consideran las características de las restricciones de tráfico ofrecido y capacidad de la red cuando toman decisiones de encaminamiento. Esto da lugar a que subconjuntos de recursos de la red estén congestionados, mientras que otros recursos en caminos alternativos permanecen en niveles bajos de utilización. Este tipo de problema de congestión es un síntoma de una asignación de recursos mala, y es un tema que la Ingeniería de Tráfico intenta mejorar.

MPLS permite que se introduzcan en las redes IP sofisticadas técnicas de control de encaminamiento. Un camino de etiqueta conmutado (*Label-Switched Path, LSP*) se determina en el nodo origen, y el control de conexión en origen permite establecer rutas explícitas independientes de la ruta

establecida por el algoritmo de encaminamiento IP. Una vez que se ha determinado la ruta, se usa un protocolo de señalización para instalar el LSP (también llamado Túnel MPLS). A través del LSP y sobre el modelo de encaminamiento IP, se superpone una capacidad similar a la de conmutación de circuitos. MPLS simplifica, de esta forma, la arquitectura y el diseño de la red.

Los túneles LSP permiten optimizar el rendimiento de una red de varias formas. Por ejemplo, si se detectan problemas de congestión causados por decisiones de encaminamiento subóptimas, los túneles LSP se pueden re-encaminar para solucionar el problema. Además, entre dos nodos se pueden crear múltiples túneles LSP, y el tráfico entre ellos se puede dividir entre los túneles de acuerdo a alguna política local. Estos túneles también se pueden parametrizar, y los recursos de la red se pueden asignar a ellos en función de dichos parámetros, es decir, los túneles LSP permiten la introducción de opciones de supervivencia flexibles. Además, las estadísticas obtenidas en estos túneles se pueden usar para construir una matriz de tráfico.

La gestión del camino es el componente de MPLS que se encarga de todos los aspectos relacionados con la selección de rutas explícitas, y de la instanciación y mantenimiento de los túneles LSP. La gestión del camino está formado por tres funciones primarias: selección del camino, localización del camino y mantenimiento del camino. La función de selección del camino especifica la ruta explícita para un túnel LSP en el nodo origen del túnel. Una ruta explícita se puede representar como una secuencia de saltos o una secuencia de nodos, y se pueden definir administrativamente o se pueden calcular automáticamente por una entidad de encaminamiento basándose en unas restricciones específicas. Esta entidad se usa para calcular las rutas que satisfacen el conjunto de requisitos, sujetos a las restricciones impuestas por la red y las políticas administrativas; de esta forma se reduce el nivel de intervención manual involucrada en la Ingeniería de Tráfico. El segundo componente de la gestión del camino es la función de localización del camino. Esto se utiliza para instanciar los túneles LSP usando un protocolo de señalización, que también se utiliza como un protocolo de distribución de etiqueta. Actualmente hay definidos varios protocolos de señalización MPLS, entre ellos se encuentran: RSVP (*Resource Reservation Protocol*) y CR-LDP (*Constraint-based Routed LDP*). El tercer componente de la gestión del camino es la función de mantenimiento del camino, que sostiene y termina túneles LSP ya establecidos.

A continuación, una vez que el túnel se ha establecido se le debe asignar el tráfico. La asignación de tráfico está formada por una función de partición y una función de asignación. La función de partición se encarga de dividir el tráfico entrante en función de algún principio específico de división. La función de asignación asigna el tráfico dividido a un túnel LSP establecido, de acuerdo a algún principio de asignación.

La función de distribución de la información del estado de la red se encarga de la distribución de la información relevante del estado de la topología a través del dominio MPLS. Esto se desarrolló con la extensión de los protocolos IGP (*Interior Gateway Protocol*) tradicionales para que propagaran información adicional acerca del estado de la red. La información adicional distribuida incluye ancho de banda máximo en el enlace, ancho de banda reservado por prioridad de clase, y atributos de la clase de recursos. Esta información es utilizada por el método de encaminamiento para seleccionar rutas factibles para los túneles LSP.

RSR (*Random Shaped Routing*) es una estrategia de encaminamiento dinámico que se puede aplicar en MPLS. RSR realiza en las redes de datos la misma tarea que DAR hace en redes telefónicas, y el funcionamiento del algoritmo es similar. La decisión de encaminamiento consiste generalmente en elegir la mejor ruta, pero también es necesario repartir regularmente la carga entre todos los caminos equivalentes. Esto se puede lograr seleccionando aleatoriamente el camino con una probabilidad proporcional a la calidad estimada del camino. Por lo tanto, al encaminar aleatoriamente M caminos con igual longitud y ancho de banda, todos reciben la misma cantidad de tráfico. Otro aspecto importante es que con el encaminamiento aleatorio la carga es repartida entre todos los caminos disponibles; el protocolo de encaminamiento incluso tiene en cuenta los cambios en la carga del enlace causados por la presencia de un nuevo tráfico entrante. Esto es muy valioso cuando un gran número de conexiones entran en la red al mismo tiempo, que ocurre, por ejemplo, cuando un componente de la red se cae y el tráfico

que estaba usando ese componente es distribuido en la red. Otras variaciones del DAR propuestas envían el tráfico a las rutas alternativas en una relación inversamente proporcional a la carga que ya han transportado, pero sólo si la ruta directa está saturada. La construcción del camino en RSR sigue un principio muy simple, ya que el camino se construye incrementalmente, comenzando en el nodo origen de las peticiones. A continuación, una vez que se haya seleccionado uno de los vecinos, la misma operación se repetirá de nodo en nodo.

De esta forma se puede comprobar que los antiguos algoritmos de encaminamiento dinámico en redes de conmutación de circuitos se pueden adaptar para utilizarlos en la fase de selección del camino de MPLS.

6.5. Simulación

Con el objetivo de comprobar el comportamiento que tienen los métodos de encaminamiento dinámico de conmutación de circuitos en redes de conmutación de paquetes, se ha programado un simulador en lenguaje C. El escenario que se pretende simular es similar al presentado en los capítulos anteriores, con la salvedad de que ahora las comunicaciones de voz se establecen mediante voz sobre IP (VoIP). Se ha elegido esta especificación por dos razones fundamentales. La primera de ellas es que las redes IP se han impuesto en los últimos años al resto de redes de conmutación de paquetes, y se presentan como el futuro más prometedor dentro de este tipo de redes. La segunda de ellas es que mediante la simulación de comunicaciones de voz en un escenario similar a los anteriores podremos comparar el comportamiento del encaminamiento dinámico en conmutación de paquetes, respecto al obtenido en conmutación de circuitos.

Se ha elegido la especificación G.729 de VoIP para la simulación. VoIP se encapsula en la parte superior del RTP, que se encapsula a su vez en la parte superior del UDP. Por tanto, VoIP es transportado con una cabecera de paquete RTP/UDP/IP. En dicha especificación, el campo de datos tiene un tamaño por defecto de 20 bytes. A este tamaño hay que añadir la longitud de la cabecera RTP (12 bytes), de la cabecera UDP (8 bytes) y de la cabecera IP (20 bytes). Vamos a suponer que en el nivel de enlace de datos el protocolo utilizado es Ethernet, por lo que también hay que añadir el tamaño de la cabecera utilizada por dicho protocolo, que por defecto es de 26 bytes. Sumándolo todo obtenemos un tamaño de paquete de 86 bytes (688 bits). En la simulación supondremos que los paquetes son de longitud fija, por lo que el tamaño de todos ellos será igual al mencionado.

Aplicación de voz
Códec de voz
RTP
UDP
IP
Ethernet

Figura 6.1 – Protocolos implicados en la transmisión de VoIP.

Nuevamente, la situación que se va a simular es la de una red simétrica totalmente conectada de 5 nodos, similar a las anteriores. En este caso vamos a suponer que los enlaces son unidireccionales y que disponen de una capacidad de 640 kbps. La elección de esta capacidad no es arbitraria sino que viene motivada por el objetivo, anteriormente comentado, de poder comparar los resultados obtenidos de esta simulación con los que se obtuvieron para conmutación de circuitos. Supondremos que en un mismo enlace puede haber 10 comunicaciones simultáneas, lo que en conmutación de circuitos se corresponde con enlaces de 10 circuitos, y ya que, como se verá a continuación, la voz se genera a una velocidad de 64 kbps, obtenemos una capacidad por enlace de 640 kbps. De esta manera compararemos los resultados obtenidos en esta simulación con los resultados obtenidos en conmutación de circuitos para enlaces de 10 circuitos.

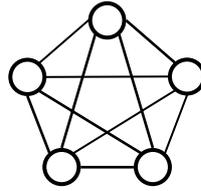


Figura 6.2 – Red simétrica totalmente conectada de cinco nodos.

El teorema de Nyquist afirma que si se muestrea una señal analógica a una velocidad dos veces superior a la frecuencia de interés más alta, se puede reconstruir de nuevo de manera exacta esa señal en su forma analógica. Como la mayoría del contenido de voz está por debajo de 4.000 Hz (4 kHz), se requiere una velocidad de muestreo de 8.000 veces por segundo (125 μ s entre muestras). En telefonía cada muestra tiene 8 bits, lo que nos da una velocidad de 64 kbps (8 bits x 8.000 Hz = 64.000 bps).

Sin embargo, una de las peculiaridades de las comunicaciones de voz sobre IP es que se permite la utilización de códecs de compresión, de forma que se reducen las necesidades de ancho de banda para este tipo de transmisiones. Un método de compresión utilizado a menudo es la modulación por impulsos codificados diferencial y adaptable (*ADPCM, Adaptive Differential Pulse Code Modulation*). Un ejemplo de utilización común de la ADPCM es la ITU-T G.726, que codifica utilizando muestras de 4 bits, lo que da una velocidad de transmisión de 32 kbps. A diferencia de la PCM (*Pulse Code Modulation*), los 4 bits no codifican directamente la amplitud de la voz, sino que codifican las diferencias de la amplitud entre dos muestras consecutivas, así como la velocidad de cambio de esa amplitud, empleando alguna predicción lineal rudimentaria.

Cuando se utiliza G.729, que es la especificación que vamos a simular, el DSP (Procesador Digital de Señal) genera una muestra de voz de 10 bytes cada 10 ms, ya que proporciona una velocidad de compresión de 8 kbps. Dos de esas muestras de voz (ambas con 10 ms de retraso) se colocan dentro de un paquete. El retraso de paquete es, por tanto, de 20 ms. Los fabricantes pueden decidir cuantas muestras de voz quieren enviar en un paquete. Como G.729 utiliza muestras de voz de 10 ms, cada incremento en las muestras por trama aumenta el retraso en 10 ms.

Por otra parte, puesto que los paquetes generados van a ser de tamaño constante e igual a 86 bytes, y la capacidad establecida para los enlaces es de 640 kbps, obtenemos que el tiempo de servicio de los paquetes, también constante, es de 1,075 ms.

Existe una regla que dice que el tamaño en bits de cada uno de los paquetes generados en VoIP debe ser menor al producto del ancho de banda de la línea por el 80% del tiempo de cadencia de generación del paquete. En nuestro caso el tamaño máximo de paquete permitido es de $640.000 \text{ bps} \cdot 0,8 \cdot 0,02 \text{ s} = 10.240 \text{ bits}$ que, como se puede ver, es superior a los 688 bits que vamos a utilizar como tamaño de paquete.

A continuación se van a describir dos fenómenos que aparecen en las transmisiones de voz sobre IP, como son el retraso inicial y el *jitter*. El retraso o latencia en VoIP se caracteriza por el tiempo que tarda la voz en salir del terminal del que está hablando hasta llegar al terminal del que está escuchando. Existen tres tipos de retraso que son inherentes a las redes de telefonía actuales: retraso de serialización, retraso de propagación y retraso de paquetización. El retraso de serialización es la cantidad de tiempo que se tarda en colocar un bit o byte en un interfaz. Su influencia en el retraso es relativamente pequeña. El retraso de propagación es causado por la velocidad de la luz en la fibra óptica o de la electricidad en las redes basadas en cobre. La luz viaja a través del vacío a una velocidad de 300.000 kilómetros por segundo y los electrones viajan a través del cobre a unos 200.000 kilómetros por segundo. Una red de cobre alrededor del mundo (21.000 kilómetros) induce un retraso de sentido único de unos 70 milisegundos, que puede provocar una degradación apreciable de la voz. El retraso de paquetización es provocado por los distintos dispositivos que envían tramas a través de la red. Los retrasos de paquetización pueden tener impacto en las redes telefónicas tradicionales, pero esos retrasos son un problema mayor en los entornos

de paquetes. El DSP genera una muestra de voz cada 10 ms cuando se utiliza G.729, como se ha indicado anteriormente. Dos de esas muestras de voz (ambas con 10 ms de retraso) se colocan dentro de un paquete. El retraso de paquetización es, por tanto, de 20 ms.

Una red basada en paquetes sufre un retraso en la gestión de las colas. Este tipo de retrasos ocurre cuando se envían más paquetes que los que la interfaz puede manejar en un intervalo de tiempo dado. La recomendación G.114 de la ITU-T especifica que para una buena calidad de voz no debe darse un retraso inicial mayor de 150 ms.

En las redes basadas en conmutación de paquetes los distintos flujos de información deben atravesar una serie de colas. El comportamiento de estas colas, es decir, el retardo que se experimenta en la gestión de las colas de salida de los nodos viene controlado por distribuciones estadísticas que representan variables aleatorias. La variación del retardo o *jitter* mide la variación que presentan estas variables aleatorias. Éste es el parámetro fundamental para controlar la correcta recepción de la voz. Las recomendaciones dicen que el retraso en la gestión de colas que provoca el *jitter* debe ser menor a 10 ms. El *jitter* típico en G.729 es de 6 ms.

Con el objetivo de simplificar la simulación, asumiremos que el valor de los dos parámetros anteriores, retraso inicial y *jitter*, debe ser menor a 10 ms.

En el simulador, con el objetivo de mantener los dos parámetros anteriores debajo del umbral establecido en el párrafo anterior, en el momento de establecer la conexión se van a comprobar el estado de todas las colas que deben atravesar los paquetes de dicha conexión. En el caso de que el número de paquetes en cada una de las colas se encuentre por debajo de un determinado umbral K preestablecido se permitirá el establecimiento de la conexión. En caso contrario, se hará que la llamada se desborde hacia el camino que determine el algoritmo de encaminamiento dinámico, o se bloqueará la conexión, según corresponda.

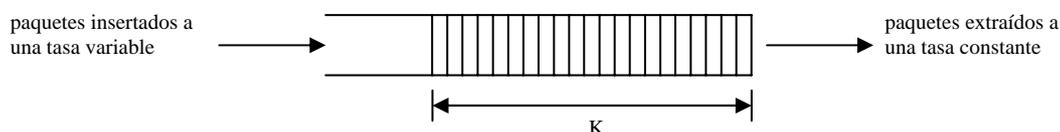


Figura 6.3 – Modelización de las colas para controlar el retraso inicial y el *jitter*.

El cálculo del umbral anterior se realiza teniendo en cuenta que el tiempo de servicio de los K paquetes permitidos en cada una de las colas atravesadas debe ser menor a los 10 ms establecidos anteriormente y que, como máximo, los paquetes de una misma conexión deben atravesar dos colas. El tiempo de servicio de los paquetes, tal y como se dijo anteriormente, es constante e igual a 1,075 ms, por lo tanto el producto $2 \cdot K \cdot 1,075$ debe ser inferior a los 10 ms. Uno de los valores de K utilizados en la simulación y que cumplen la restricción anterior es de $K = 4$. Por otra parte, ya que, como se dijo anteriormente, cuando se utiliza la codificación G.729 la variación del retardo toma típicamente el valor de 6 ms, las simulaciones típicas, que utilizaremos para comparar los resultados obtenidos con el encaminamiento dinámico, se van a realizar tomando un valor de umbral $K = 3$.

Decir también, que el algoritmo de encaminamiento dinámico utilizado en el simulador es el más general de todos. Es decir, aquel en el que la elección de la ruta alternativa se realiza de forma totalmente aleatoria, el algoritmo *aleatorio puro*. Dado que la red simulada está formada por 5 nodos ($N=5$), se permitirá que en la búsqueda de la ruta secundaria se compruebe el estado de cada una de las 3 posibles rutas alternativas ($M=3$), mientras no se llegue al establecimiento de la conexión con éxito.

Por otra parte, también hay que señalar que en el nuevo simulador, tanto el tiempo de duración de las llamadas como el tiempo entre dos llamadas consecutivas siguen estando distribuidos exponencialmente, representando una distribución de tráfico de Poisson. Algunas de las diferencias en la programación de este simulador respecto al de los capítulos anteriores vienen provocadas por la propia

naturaleza de la arquitectura que es objeto de simulación. Así, se puede apreciar que encada uno de los 4 enlaces de salida que posee un nodo debe haber una cola con el objeto de poder almacenar los paquetes que no se pueden servir en el momento en el que intentan hacerlo. Estas colas utilizan prioridades (802.1p) con el objetivo de asignar la mayor capacidad posible las comunicaciones de voz frente a las comunicaciones de datos, de manera que las llamadas de voz se encuentran toda la capacidad de los enlaces como si sólo se estableciera este tipo de comunicación, dejando el ancho de banda restante para las comunicaciones de datos. Por otra parte, durante el tiempo de duración de la conexión, distribuido exponencialmente, la llamada generará paquetes de voz en intervalos de tiempo fijos, debido a que los paquetes son de tamaño fijo, como ya se comentó anteriormente.

6.5.1. Resultados de simulación

Las *prestaciones de la red* se van a establecer, de nuevo, como la probabilidad de pérdida en llamadas, es decir, la probabilidad de que una llamada encuentre la ruta directa y todas las secundarias por encima del umbral K preestablecido. Como ya se dijo en el capítulo 2, a este parámetro también se le conoce como *grado de servicio* (GoS). Por lo tanto, uno de los resultados extraídos de la simulación y que será representado gráficamente será el GoS.

Por otra parte, también se va a presentar como resultado de la simulación la *utilización* del enlace por parte de la transmisión de VoIP. Dicha parámetro se calcula como el cociente entre el tiempo en que el enlace está ocupado en una comunicación VoIP determinada y el tiempo total de simulación. Este parámetro tiene una gran importancia porque hay que tener en cuenta que el ancho de banda disponible por el usuario se debe compartir entre la transmisión de VoIP y otras posibles comunicaciones de datos realizadas de forma simultánea.

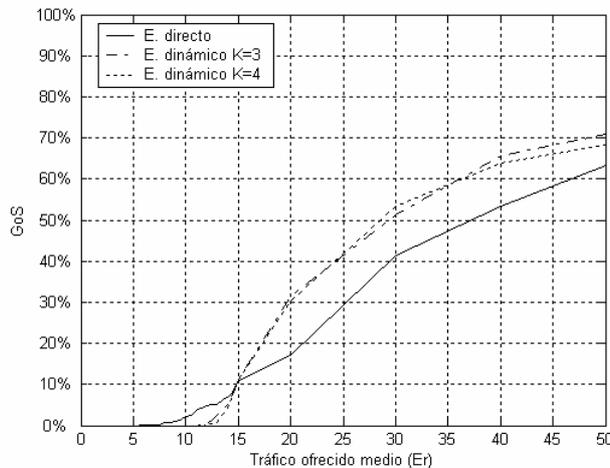


Figura 6.4 – Prestaciones del encaminamiento directo y del dinámico, para dos valores de K .

En la figura 6.4 se representan conjuntamente los GoSs que se obtienen con el encaminamiento directo y con el dinámico para un amplio margen de tráfico. Los valores representados para el encaminamiento directo son los correspondientes a un valor del parámetro K igual a 3 que, como ya se dijo anteriormente, representa el valor típico del *jitter* permitido al utilizar la especificación G.729. Se puede comprobar que se obtienen una resultados equivalentes a los obtenidos en capítulos anteriores para algoritmos en encaminamiento dinámico en redes de conmutación de circuitos. Es decir, los algoritmos que contemplan el encaminamiento alternativo presentan mejores prestaciones para tráficos bajos, donde las rutas alternativas solventan el problema transitorio de bloqueo en el enlace directo, y peores para cargas altas, donde el uso de dos enlaces es más frecuente y provoca un aumento en el número de llamadas rechazadas. Debemos recordar que uno de los parámetros característicos de los servicios en redes de conmutación de circuitos es el GoS máximo en la hora cargada y que este valor no suele superar el 2 %.

Si comparamos los resultados mostrados en la figura 6.4, con los obtenidos en redes de conmutación de circuitos, podemos apreciar a simple vista dos diferencias fundamentales. La primera de ellas es que el umbral para el cual se producen mejoras en el encaminamiento dinámico respecto al directo es mayor al que se obtenía para redes de conmutación de circuitos. La segunda diferencia es que puede comprobarse que la separación entre las curvas de encaminamiento directo y dinámico es mayor que la que se tenía para redes de conmutación de circuitos.

Los resultados de la simulación dentro del margen significativo de valores, se muestran en las siguientes figuras. En primer lugar, en la figura 6.5 podemos comparar más en profundidad los resultados obtenidos en función del parámetro K . Así, observamos que para un mayor valor de este parámetro el umbral para el cual empiezan a aparecer pérdidas también se incrementa, ya que para un valor de $K=3$ el umbral anterior se muestra en 11 Erlangs aproximadamente, mientras que para un valor de $K=4$ se sitúa en un valor de tráfico ofrecido aproximado de 12 Erlangs. También observamos que el punto de cruce de ambas curvas con la del encaminamiento directo se produce para un valor de tráfico ofrecido muy cercano, entorno a los 14,5 Erlangs aproximadamente.

Las dos circunstancias anteriores, es decir, que el umbral a partir del cual aparecen pérdidas sea mayor para $K=4$, y que el cruce de las curvas se produzca para un mismo valor de tráfico ofrecido, hace que por necesidad, la pendiente de la curva con $K=4$ sea superior a la curva con $K=3$.

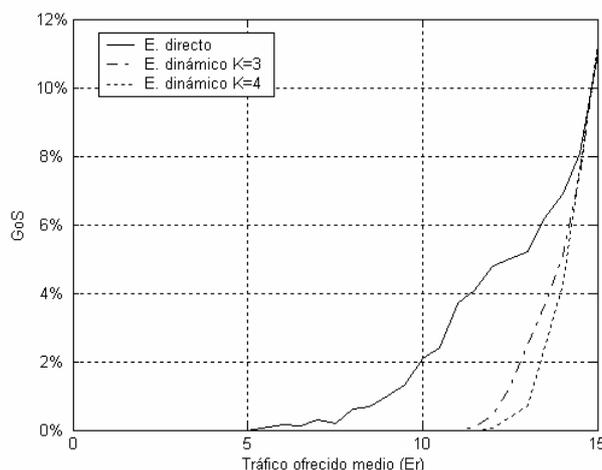


Figura 6.5 – GoS de la red para diferentes valores medios de tráfico ofrecido entre cada par de nodos.

A continuación se va a realizar un estudio, dentro del margen significativo de valores, de las figuras relativas a las prestaciones en GoS, similar al realizado en capítulos anteriores. Se muestran las curvas de evolución de un parámetro frente al otro para encaminamiento directo y encaminamiento dinámico con un valor de $K=3$ en ambos casos, la ganancia absoluta y la ganancia relativa, calculadas según (3.1a) y (3.1b).

Debido a que sólo se representan dentro del margen más significativo de valores, se comprueban las mejores prestaciones del encaminamiento dinámico. Sin embargo, las curvas de prestaciones (figura 6.6) se van acercando a medida que se aumenta el tráfico, anunciando la situación mostrada en la figura 6.4. También se comprueba que el encaminamiento dinámico es capaz de cursar una gran cantidad de tráfico sin presentar pérdidas. Podría fijarse un umbral aproximado de 11 Erlangs, frente a los 5 Erlangs para el encaminamiento directo.

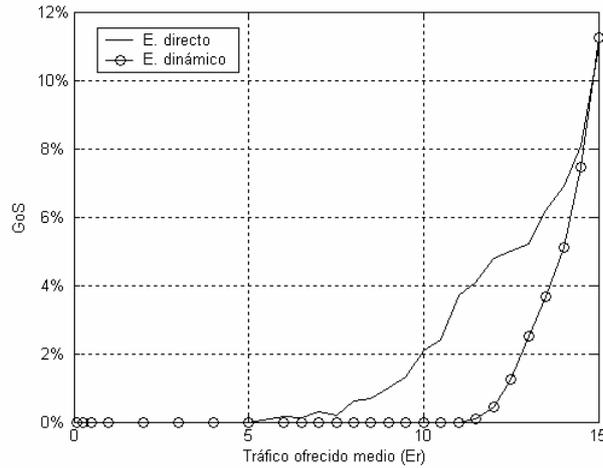


Figura 6.6 – GoS de la red para diferentes valores medios de tráfico ofrecido entre cada par de nodos.

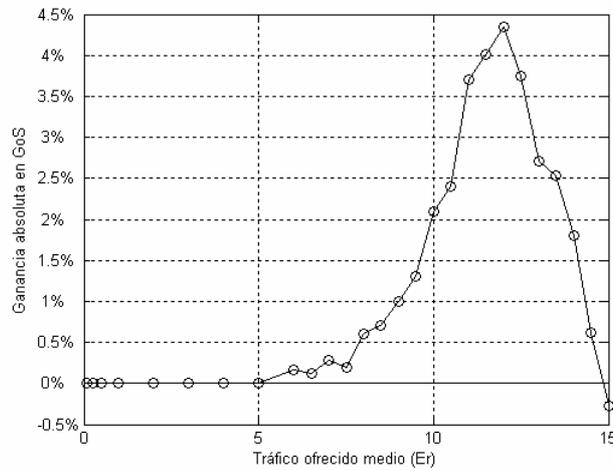


Figura 6.7 – Ganancia absoluta en GoS.

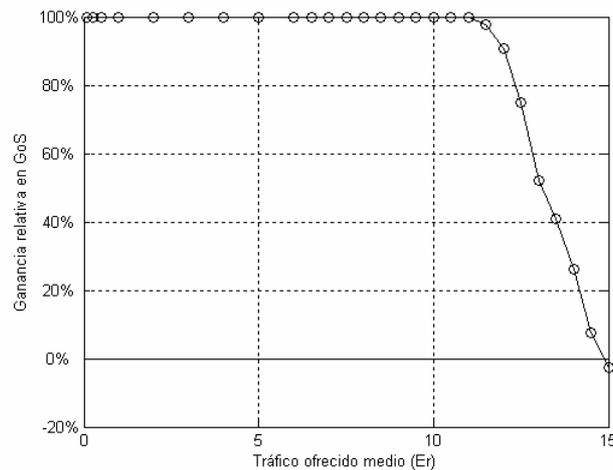


Figura 6.8 – Ganancia relativa en GoS.

La curva de ganancia absoluta se corresponde con la del directo hasta los 11 Erlangs aproximadamente. Más allá de ese valor, el encaminamiento dinámico crece con una mayor pendiente por lo que su GoS llega a alcanzar al directo, cayendo, por lo tanto la ganancia absoluta hasta hacerse

negativa. El máximo se alcanza aproximadamente para los 12 Erlangs, donde se mejora el GoS en un 4,3465 %.

La ganancia relativa, mostrada en la figura 6.8 muestra que los máximos se obtienen para valores muy bajos de tráfico ofrecido, debido a la normalización que se realiza dividiendo por las prestaciones del encaminamiento directo, valores menores para menores tráficos.

Dado que ambas curvas (encaminamiento directo y encaminamiento dinámico) son coincidentes para tráficos muy bajos y se cruzan a partir de un determinado valor, la diferencia entre ambas presentará un máximo en el interior del intervalo definido. Los valores máximos de ganancia y los puntos en los que se producen se muestran de forma aproximada en la tabla 6.1.

G absoluta		
Prest.	Máximo en	G
GoS	A = 12 Er	4,3465%

G relativa		
Prest.	Máximo en	G
GoS	A < 6 Er	100,00%

Tabla 6.1 – Ganancias absoluta y relativa máximas para GoS.

A continuación se van a analizar los resultados obtenidos en la simulación para el parámetro de *utilización*. Observando la figura 6.9 se pueden resaltar a simple vista varias circunstancias importantes. En primer lugar, se puede apreciar como este parámetro presenta una evolución lineal similar, tanto para el encaminamiento directo, como para el encaminamiento dinámico con valores de $K=3$ y $K=4$, hasta un valor de tráfico ofrecido de 12 Erlangs aproximadamente. A partir de ese valor la evolución para los dos tipos de encaminamiento sigue una distribución distinta, siendo superior la utilización cuando se utilizan algoritmos de encaminamiento dinámico que cuando se utiliza únicamente el encaminamiento directo. Además, se puede apreciar que entorno a un valor de tráfico ofrecido de 15 Erlangs aproximadamente se produce un punto de inflexión, relajándose el ritmo de crecimiento de la utilización y tendiendo de forma asintótica hacia el límite de utilización del 100%. Este punto es muy curioso, puesto que coincide con el punto en el que se cruzan las curvas de prestaciones en GoS para métodos de encaminamiento dinámico y para el encaminamiento directo. También se puede observar que entre 15 y 35 Erlangs aproximadamente, las prestaciones obtenidas en el encaminamiento dinámico con $K=4$ son peores a las obtenidas con $K=3$. Sin embargo, el estudio en esta zona no tiene ninguna importancia, al estar fuera del margen de valores en que el GoS es suficientemente pequeño, menor al 2%.

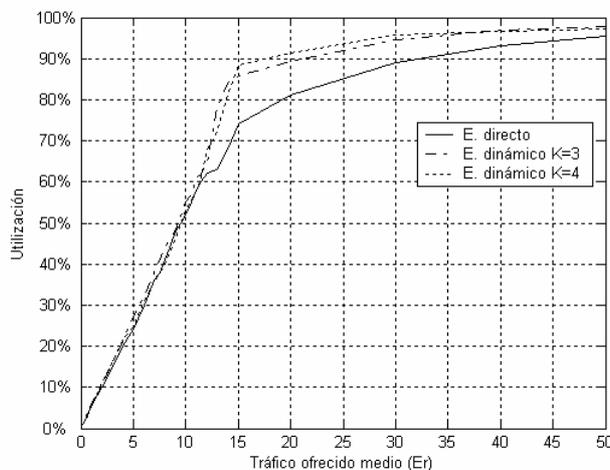


Figura 6.9 – Utilización del encaminamiento directo y del dinámico, para dos valores de K.

En la figura 6.10 aparece la utilización para aquellos valores de tráfico ofrecido que producen valores de GoS suficientemente pequeños. Se puede apreciar que para las distintas simulaciones la evolución de este parámetro es bastante similar. Sin embargo, se observa que, de forma más o menos general, el valor de la utilización obtenido en el encaminamiento dinámico con $K=3$ es peor al obtenido con $K=4$, de forma que de nuevo aquí las prestaciones mejoran al aumentar el valor del parámetro K .

Sin embargo, se observan una serie de inconvenientes en la evolución de la utilización. El hecho de que la pendiente de estas curvas sea tan pronunciada hace que para valores de tráfico ofrecidos no muy altos se alcancen valores de utilización elevados, aunque durante un amplio margen de tráfico ofrecido, los métodos de encaminamiento dinámico no empeoran de forma sustancial los valores de utilización obtenidos con encaminamiento directo. Es a partir de un valor de 12 Erlangs aproximadamente cuando la utilización de métodos de encaminamiento dinámico incrementa la utilización del enlace.

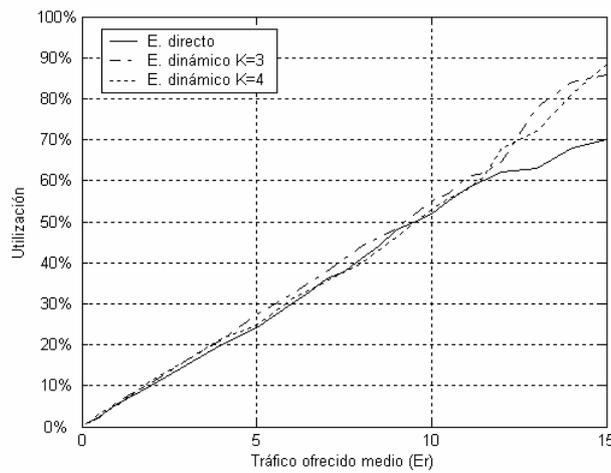


Figura 6.10 – Utilización de la red para diferentes valores medios de tráfico ofrecido entre cada par de nodos.

En la tabla 6.2 se observan los valores de tráfico ofrecido y utilización obtenidos para un GoS máximo del 1%, para encaminamiento directo y dinámico con un valor de $K=3$. Se puede apreciar como el encaminamiento directo alcanza el mencionado valor de GoS para un valor de tráfico ofrecido de 9 Erlangs, que como se puede observar en la figura 6.10, se encuentra en la zona en la que la evolución de la utilización es similar para los encaminamiento directo y dinámico, correspondiéndole una utilización del 48 %. Sin embargo, para el encaminamiento dinámico se obtiene el 1% en GoS para un tráfico ofrecido de 12,34 Erlangs, que se encuentra en la zona en que empeora la utilización al utilizar encaminamiento dinámico, correspondiéndole un valor de 69,4%, extremadamente alto.

GoS (%)	Ao dir. (Er)	Ao din. (Er)	U dir. (%)	U din. (%)
1	9	12,34	48	69,4

Tabla 6.2 – Tráficos ofrecidos y utilizations para GoS máximo.

En el análisis de este parámetro debemos caer en un aspecto importante. Y es que, si comparamos los resultados de la evolución del GoS con los obtenidos para conmutación de circuitos, observamos que el valor del 1% se obtiene para valores de tráfico ofrecido 5,75 Erlangs para el DAR, y 6,25 Erlangs para el LLR. Si nos vamos con estos valores de tráfico ofrecido a la figura 6.10, que representa la evolución para conmutación de paquetes, observamos que en ambos casos nos encontramos en la zona en la que la utilización obtenida no empeora de forma sustancial la que le corresponde al encaminamiento directo. Por lo que podemos concluir que la utilización de algoritmos de encaminamiento dinámico en redes de conmutación de paquetes no mejora la utilización obtenida con respecto a la utilización única del encaminamiento directo, pero tampoco la empeora.

Llegados a este punto hay que hacer varios razonamientos. En primer lugar, como ya se observó anteriormente, los métodos de encaminamiento dinámico no introducen ningún tipo de mejoras respecto

al encaminamiento directo en lo que se refiere a la utilización del enlace. En segundo lugar, en lo que se refiere a los resultados obtenidos para el grado de servicio, observamos que el comportamiento de los métodos de encaminamiento dinámico en redes de conmutación de paquetes es similar al obtenido en redes de conmutación de circuitos.

Sin embargo, hay que centrarse en otro aspecto. Si fijamos una cota máxima de GoS del 1%, como hicimos en la tabla 6.2, utilizando únicamente el encaminamiento directo podemos ofrecer al enlace un tráfico de hasta 9 Erlangs. Este tráfico ofrecido ya es superior al que podemos ofrecer en redes de conmutación de circuitos que utilizan métodos de encaminamiento dinámico que, como se comentó en párrafos anteriores, toman valores de 5,75 y 6,25 Erlangs para el DAR y el LLR, respectivamente. Además, hay que resaltar que las simulaciones realizadas asumían enlaces de 640 kbps, que es una capacidad enormemente superada en los enlaces troncales de Internet, que llegan hasta valores de 10 Gbps. Por lo tanto, si asumiéramos capacidades para los enlaces con valores más reales, obtendríamos que utilizando únicamente el encaminamiento directo podríamos ofrecer elevados valores de tráfico a la red hasta llegar a un valor de GoS máximo del 1%.

Por lo tanto, se puede concluir que el uso de los métodos de encaminamiento dinámico en redes de conmutación de paquetes con el objetivo de disminuir el GoS de la red, permitiendo el uso de rutas alternativas por donde cursar las llamadas, que es el que tenía en las redes de conmutación de circuitos, no es necesario de ninguna forma, pues las prestaciones que se obtienen con el encaminamiento directo ya son excelentes.

6.6. Conclusiones

En este capítulo se ha comprobado de forma práctica que el uso de los métodos de encaminamiento dinámico en redes de conmutación de paquetes (concretamente en VoIP), con el objetivo de disminuir el GoS de la red, que es el que tenían en redes de conmutación de circuitos, es totalmente innecesario, pues las prestaciones que se obtienen con el encaminamiento directo ya son muy buenas.

Sin embargo, se ha comprobado que algunas de las características de estos métodos se han adaptado e implementado en redes de conmutación de paquetes existentes.

Así, en las redes ATM, con el objetivo de optimizar el ancho de banda disponible para los distintos circuitos virtuales, se han extendido conceptos como la reserva de troncales y el encaminamiento alternativo controlado. Estas aproximaciones están incluidas en el protocolo PNNI del Forum ATM.

Por otra parte, en las redes IP podemos ver métodos de encaminamiento dinámico utilizados de forma conjunto con métodos de reserva de ancho de banda y aplicados a la gestión de las tablas de encaminamiento. Además, se ha comprobado que logran un rendimiento muy bueno en la gestión de los caminos MPLS.

Con MPLS el gran desafío que se pretende conseguir es el de poder transportar diferentes tipos de información entre los usuarios, asegurando una cierta calidad de servicio. Para ello, esta tecnología permite introducir sofisticadas técnicas de control de encaminamiento en redes IP. Los túneles LSP permiten la detección de problemas de congestión causados por decisiones de encaminamiento subóptimas y el re-encaminamiento para solucionar el problema. En este tipo de tecnologías existe una aplicación directa de los métodos de encaminamiento dinámico para realizar la selección del camino en el que posteriormente se instanciará el túnel LSP. De esta forma, se ha introducido RSR, que es una estrategia de encaminamiento dinámico aplicable en MPLS, y que se basa en el algoritmo DAR, estudiado en el capítulo 3.

Capítulo 7

Conclusiones y líneas futuras

7.1. Conclusiones

En el primer capítulo de este proyecto se plantearon diferentes objetivos a partir de la idea de evaluar, de forma teórica y práctica, las prestaciones de los métodos de *encaminamiento dinámico no jerárquico* en redes de conmutación de circuitos.

En el presente capítulo se analiza si dichos objetivos han sido cubiertos, extrayendo conclusiones sobre los estudios teóricos, la comparación entre los distintos tipos de métodos de encaminamiento dinámico y entre los resultados obtenidos en las simulaciones de DAR y LLR, y la aplicación que estos métodos pueden tener en redes de conmutación de paquetes.

En la fase inicial del proyecto se ha realizado un estudio teórico sobre el encaminamiento dinámico en redes no jerárquicas, y sobre el encaminamiento cuando se admite la reserva de troncales para uso exclusivo del tráfico directo entre cada par de nodos. Hay que indicar que el objetivo de estos métodos es reducir la probabilidad de pérdida en llamada, es decir, la probabilidad de que una llamada encuentre la ruta directa y todas las alternativas bloqueadas. Este parámetro también recibe el nombre de grado de servicio.

En este estudio teórico se ha demostrado que bajo ciertas condiciones aparecen márgenes de inestabilidad, es decir, se ha demostrado la existencia de más de una solución en la probabilidad de pérdida para un rango de tráfico ofrecido. Esta inestabilidad, que presenta toda red con encaminamiento dinámico no jerárquico, no supondría ningún problema si dichas soluciones permaneciesen dentro de valores tolerables de la probabilidad de pérdida. Sin embargo, en el estado de alta probabilidad de pérdidas, ésta es mucho mayor que la que se obtiene con encaminamiento directo. La solución a este problema pasa por la utilización de la reserva de troncales. Se ha demostrado que la reserva de un número pequeño de troncales ya soluciona el problema de la inestabilidad.

A continuación se han evaluado de forma teórica y práctica, mediante simulación, dos algoritmos concretos, como son el DAR y LLR. De la comparación entre las prestaciones obtenidas entre estos dos algoritmos podemos extraer varias conclusiones. En primer lugar, se comprueban las mejores prestaciones del LLR. Sin embargo no son muy superiores a las del DAR, y además dependen del margen de tráfico ofrecido y GoS en el que trabajamos, tal y como se observa en las curvas de prestaciones y de ganancias absolutas. Pueden haber situaciones en las que no nos importe conseguir un rendimiento ligeramente inferior, si a cambio se puede conseguir reducir los costes y los problemas de escalabilidad asociados al LLR. Por ello, hay que decir que la utilización en redes reales del DAR, a pesar de su gran simplicidad, se puede justificar de una forma muy razonada, ya que logra conseguir unos rendimientos muy aceptables.

Siguiendo con este tema, han quedado demostradas las ventajas de utilizar un método de encaminamiento distribuido frente a uno centralizado y las ventajas de los métodos adaptativos frente a aquellos que tienen un comportamiento dependiente del tiempo. También se han presentado los inconvenientes que trae consigo el requerir información global sobre el estado de la red para ejercer las funciones propias de los métodos de encaminamiento dinámico, derivadas sobre todo de los costes añadidos de una red de señalización, y de las dificultades de escalabilidad que tienen los métodos que necesitan información de los distintos nodos de la red para su funcionamiento. Por último, se ha demostrado que las reglas de selección de ruta basadas en el principio de *sticky-random*, como las que utilizan algunos métodos EDR, además de su gran simplicidad y la ausencia de necesidad de una red

señalización, logran unos rendimientos comparables con los de otros algoritmos de encaminamiento dinámico más complejos.

Por último, se ha realizado un estudio sobre la aplicación que los métodos estudiados en este Proyecto Final de Carrera pueden tener en redes de conmutación de paquetes. En primer lugar, se han presentado diversos casos de redes reales en los que se ha realizado una adaptación de diversos conceptos utilizados en redes de conmutación de circuitos. Hay que decir que en las situaciones que se han presentado no se utilizan estos métodos para conseguir los objetivos que perseguían en conmutación de circuitos, sino que se utilizan de forma conjunta con métodos de reserva de ancho de banda, aplicados a la gestión de tablas de encaminamiento o asociados a la selección del camino en MPLS. Se ha demostrado que la utilización de estos métodos con el objetivo de reducir la probabilidad de pérdida en llamadas es totalmente innecesaria, pues las prestaciones obtenidas con el encaminamiento directo ya superan enormemente las obtenidas en conmutación de circuitos.

7.2. Líneas futuras

Como ya se comentó en apartados precedentes de este Proyecto Final de Carrera, los métodos de encaminamiento dinámico en redes de conmutación de circuitos se encuentran en una etapa de recesión. Con la aparición de nuevas tecnologías como ATM e IP, que son orientadas a conmutación de paquetes, la preocupación por las redes TDM ha decrecido rápidamente. Se espera que con el desarrollo de estas nuevas redes, los actuales sistemas de encaminamiento dinámico desaparezcan progresivamente y sean reemplazados por otros nuevos sistemas implantados en protocolos IP.

Sin embargo, en este punto aparece uno de los grandes desafíos de las modernas redes de comunicaciones, que es transportar diferentes tipos de información entre los usuarios, asegurando una cierta calidad de servicio. Para poder proporcionar esta calidad de servicio se precisa trabajar, a diferencia de la red Internet, que trabaja en modo datagrama, con tráfico orientado a conexión, de forma similar a como han trabajado las redes telefónicas tradicionales. Además, también se precisa la utilización de un encaminamiento en origen. Pues bien, un caso práctico de esta situación viene dado por la utilización de DiffServ sobre MPLS.

MPLS permite la introducción en redes IP de sofisticadas técnicas de control de encaminamiento. La gestión del camino, que es uno de los componentes de MPLS, lleva acabo la selección de la ruta especificando el camino explícito para un túnel LSP en el nodo origen del túnel. Pues bien, dentro de esta función de MPLS tienen cabida estrategias de encaminamiento dinámico basadas en los tradicionales métodos de encaminamiento, con un funcionamiento similar al que tenían en las redes de conmutación de circuitos, pero con objetivos distintos. Así, ahora ya no se trata de disminuir la probabilidad de pérdida percibida por el usuario, sino de hacer que se cumplan requisitos estrictos de retardo y *jitter*. De esta forma podemos ver los métodos de encaminamiento dinámico asociados a estrategias de reserva de ancho de banda en redes IP o aplicados a la gestión de tablas de encaminamiento.

Anexo

Generación de variables aleatorias

A.1. Introducción

Uno de los parámetros que definen la red es el tráfico existente entre todo par de nodos. Este tráfico viene determinado por dos variables aleatorias: el tiempo entre peticiones de establecimiento consecutivas \tilde{t}_a y el tiempo de duración del servicio \tilde{t}_s .

A.2. Método de la transformación inversa

Dado que la mayoría de los lenguajes de programación poseen funciones de generación de variables aleatorias uniformemente distribuidas, se intentará encontrar un método general de generación de variables aleatorias con cualquier tipo de distribución a partir de una variable de distribución uniforme. Sea $F_{\tilde{x}}(x)$ la función de distribución de probabilidad (FDP) de la variable \tilde{x} , o función de densidad acumulada. Recuérdese que, como indica este nombre

$$f_{\tilde{x}}(x) = \frac{dF_{\tilde{x}}(x)}{dx} \quad (\text{A.1})$$

o, también,

$$F_{\tilde{x}}(x) = \int_0^x f_{\tilde{x}}(t) dt \quad (\text{A.2})$$

con $f_{\tilde{x}}(x)$ la función de densidad de probabilidad (fdp) de la variable \tilde{x} . Por definición de $F_{\tilde{x}}(x)$

$$F_{\tilde{x}}(x) = \Pr\{\tilde{x} < x\} \quad (\text{A.3})$$

Poniendo la variable \tilde{x} en función de la variable \tilde{u} , siendo ésta uniforme, se tiene que

$$\tilde{x} = G(\tilde{u}) \quad (\text{A.4})$$

$$F_{\tilde{x}}(x) = \Pr\{\tilde{x} < x\} = \Pr\{G(\tilde{u}) < x\} = \Pr\{\tilde{u} < G^{-1}(x)\} = G^{-1}(x) \quad (\text{A.5})$$

dado que \tilde{u} es uniforme. Por lo tanto,

$$F_{\tilde{x}}(x) = G^{-1}(x) \quad (\text{A.6})$$

$$F_{\tilde{x}}(x) = \tilde{u} \quad (\text{A.7})$$

$$\tilde{x} = F_{\tilde{x}}^{-1}(\tilde{u}) \quad (\text{A.8})$$

con \tilde{u} uniformemente distribuida entre 0 y 1, ya que éstos son los valores mínimo y máximo, respectivamente, de la FDP. Será éste un método genérico para la generación de variables aleatorias, si bien no es siempre aplicable ya que la $F_{\tilde{x}}(x)$ no se conoce de forma analítica.

Podría decirse que la inversa de la FDP deforma la distribución entre 0 y 1 para dar lugar a la FDP buscada. Este resultado es coherente con la idea intuitiva de la FDP. La generación de un número de forma uniforme entre 0 y 1 más la posterior proyección sobre x mediante la función de distribución hacen que sean más probables aquellos valores de x para los que la $F_{\tilde{x}}(x)$ tiene mayor pendiente, es decir, con mayor $f_{\tilde{x}}(x)$; en definitiva, los valores más probables.

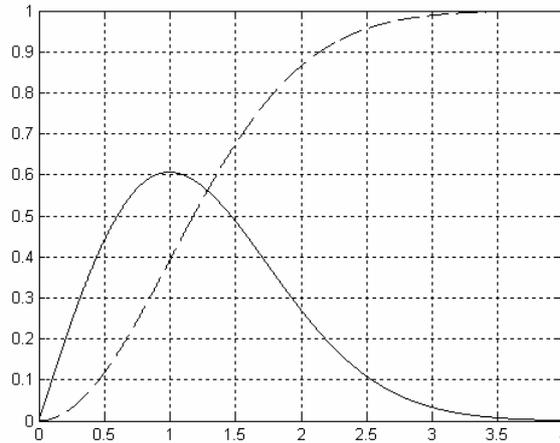


Figura A.1 – Ejemplo: FDP y fdp de Rayleigh.

A.3. Distribución determinista

Es el caso más sencillo de distribución, ya que la variable aleatoria no es tal y siempre toma un mismo valor.

$$f_{\tilde{x}}(x) = \begin{cases} 1 & x = a \\ 0 & x \neq a \end{cases} \quad (\text{A.9})$$

$$F_{\tilde{x}}(x) = u(a) \quad (\text{A.10})$$

siendo $u(a)$ la función escalón unidad. Esta función hace que la proyección de la variable uniforme generada entre 0 y 1 se concentre en un solo punto, por tanto, con probabilidad unidad.

A.4. Distribución uniforme

Sea \tilde{x} uniformemente distribuida entre a y b . En este caso,

$$f_{\tilde{x}}(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otro caso} \end{cases} \quad (\text{A.11})$$

$$F_{\tilde{x}}(x) = \begin{cases} 0 & x \leq a \\ \frac{x-a}{b-a} & a \leq x \leq b \\ 1 & x \geq b \end{cases} \quad (\text{A.12})$$

Nótese la linealidad de la función de distribución entre a y b . Por ello, al aplicar el método de la transformación inversa se traslada y escala sin deformación el intervalo $(0,1)$ hasta el (a,b) . Es decir,

$$F_{\tilde{x}} = \frac{x-a}{b-a} = u \quad (\text{A.13})$$

$$x = a + (b-a)u \quad (\text{A.14})$$

A.5. Distribución exponencial

La expresión analítica de la función de distribución exponencial es

$$f_{\tilde{x}}(x) = \begin{cases} \frac{1}{m} e^{-x/m} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (\text{A.15})$$

con $m = E\{\tilde{x}\}$. Integrando se obtiene la expresión analítica de la función de distribución buscada:

$$F_{\tilde{x}}(x) = 1 - e^{-x/m} = u = 1 - u' \quad (\text{A.16})$$

$$e^{-x/m} = u' \quad (\text{A.17})$$

$$x = -m \ln u' \quad (\text{A.18})$$

con u' también uniformemente distribuida entre 0 y 1.

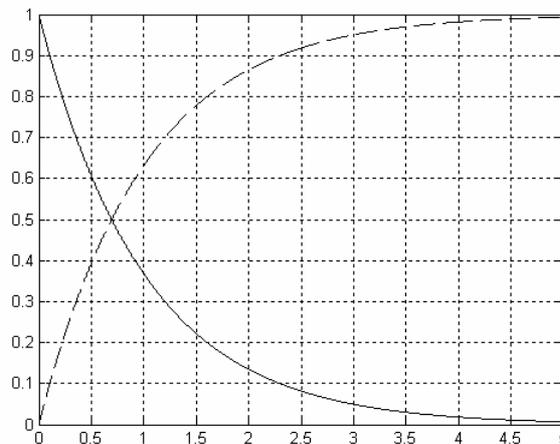


Figura A.2 – FDP y fdp exponenciales con $m=1$.

A.6. Distribución triangular

Como se muestra en la figura A.3, la función de distribución triangular puede definirse mediante el mínimo valor de la variable aleatoria (a), su moda (c) y su máximo (b). Sin embargo, puede resultar más cómodo que la función pueda expresarse en función de la media en lugar de la moda. La media se calcula como

$$E\{\tilde{x}\} = \int_{-\infty}^{\infty} xf_{\tilde{x}}(x)dx = \int_a^c xf_1(x)dx + \int_c^b xf_2(x)dx = E_1 + E_2 \quad (\text{A.19})$$

ya que

$$f_{\tilde{x}}(x) = \begin{cases} f_1(x) & a \leq x < c \\ f_2(x) & c \leq x < b \\ 0 & \text{otro caso} \end{cases} \quad (\text{A.20})$$

con

$$f_1(x) = h \frac{x-a}{c-a} \quad (\text{A.21})$$

$$f_2(x) = h \frac{b-x}{b-c} \quad (\text{A.22})$$

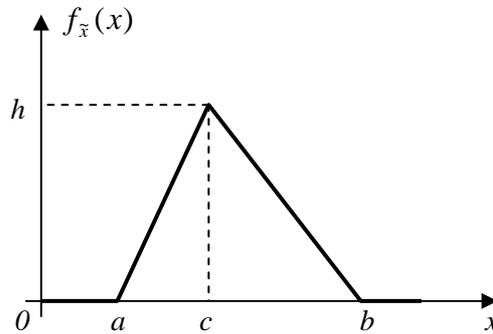


Figura A.3 – Función de distribución triangular.

Sustituyendo en la expresión del valor medio

$$E_1 = \int_a^c xh \frac{x-a}{c-a} dx = \frac{h}{c-a} \int_a^c x(x-a)dx = \frac{h}{c-a} \int_0^{c-a} (u+a)udu \quad (\text{A.23})$$

Operando, se llega a

$$E_1 = \frac{h(c-a)}{2} \left[\frac{2(c-a)}{3} + a \right] = A_1 E\{\tilde{x}_1\} \quad (\text{A.24})$$

que es el valor medio, considerando sólo el intervalo (a,c) y ponderando por el área bajo la curva en ese intervalo. Asimismo, en el intervalo (c,b),

$$E_2 = \int_c^b xh \frac{b-x}{b-c} dx = \frac{h}{b-c} \int_c^b x(b-x) dx = \frac{h}{b-c} \int_0^{b-c} (u+c)(b-c-u) du \quad (\text{A.25})$$

e, integrando,

$$E_2 = \frac{h(b-c)}{2} \left[c + \frac{(b-c)}{3} \right] = A_2 E\{\tilde{x}_2\} \quad (\text{A.26})$$

que vuelve a ser el valor medio en el intervalo, ponderado por el área correspondiente. Sumando ambos valores y operando se llega a una sencilla relación:

$$E\{\tilde{x}\} = \frac{a+b+c}{3} \quad (\text{A.27})$$

Es decir,

$$c = 3E\{\tilde{x}\} - a - b \quad (\text{A.28})$$

Obsérvese que ha de verificarse que $c \in [a, b]$, lo que implica que

$$E\{\tilde{x}\} \in \left[\frac{2a+b}{3}, \frac{a+2b}{3} \right] \quad (\text{A.29})$$

Una vez obtenida la relación anterior, los cálculos se realizarán mediante los extremos y la moda, ya que son los parámetros que permiten trabajar con una mayor comodidad, aunque las especificaciones de la variable aleatoria se realicen mediante la media. Es necesario ahora encontrar la expresión analítica de la $F_{\tilde{x}}(x)$. Se calculará por separado en los dos subintervalos.

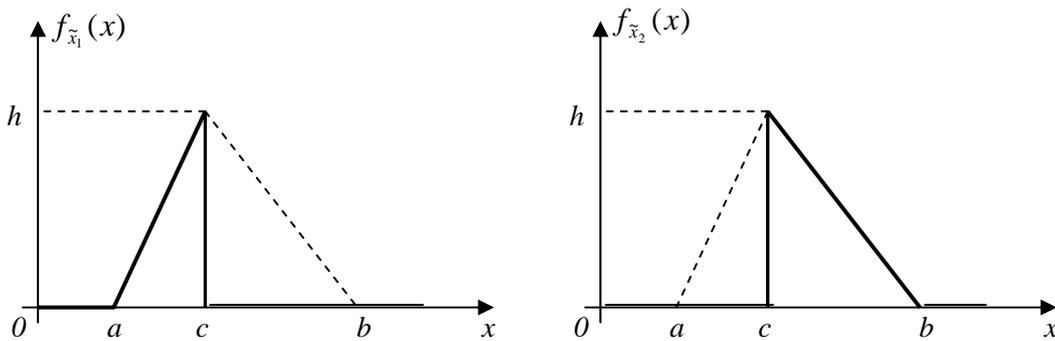


Figura A.4 – Separación de la función de distribución triangular.

En el intervalo $[a, c]$

$$F_{\tilde{x}_1}(x) = \Pr\{\tilde{x}_1 \leq x\} = \int_a^x f_{\tilde{x}_1}(y) dy = \frac{h}{c-a} \int_a^x (y-a) dy = \frac{h}{c-a} \int_0^{x-a} z dz = \frac{h}{c-a} \frac{(x-a)^2}{2} + C \quad (\text{A.30})$$

Imponiendo condiciones de contorno para calcular la constante de integración, se tiene $F_{\tilde{x}}(a) = C = 0$. También se comprueba un resultado lógico

$$F_{\tilde{x}_1}(c) = \frac{h(c-a)}{2} = \frac{c-a}{b-a} \quad (\text{A.31})$$

que es el área bajo la curva en el intervalo considerado, teniendo en cuenta que el área total bajo la fdp ha de ser unidad y expresando h , por tanto, como

$$h = \frac{2}{b-a} \quad (\text{A.32})$$

Igualando $F_{\tilde{x}_1}(x)$ a u , uniformemente distribuida en $\left(0, \frac{c-a}{b-a}\right)$ y despejando x

$$F_{\tilde{x}_1}(x) = \frac{h}{c-a} \frac{(x-a)^2}{2} = u \quad (\text{A.33})$$

$$x = a \pm \sqrt{u(c-a)(b-a)} \quad (\text{A.34})$$

Análogamente, en el intervalo $[c,b]$

$$\begin{aligned} F_{\tilde{x}_2}(x) &= \Pr\{\tilde{x}_2 \leq x\} = \int_c^x f_{\tilde{x}_2}(y) dy = \frac{h}{b-c} \int_c^x (b-y) dy = \\ &= \frac{h}{b-c} \int_0^{x-c} (b-c-z) dz = 2 \frac{x-c}{b-a} - \frac{(x-c)^2}{(b-a)(b-c)} + C' \end{aligned} \quad (\text{A.35})$$

Imponiendo continuidad en $x=c$,

$$F_{\tilde{x}_1}(c) = F_{\tilde{x}_2}(c) = C' = \frac{c-a}{b-a} \quad (\text{A.36})$$

quedando, pues,

$$F_{\tilde{x}_2}(x) = 2 \frac{x-c}{b-a} - \frac{(x-c)^2}{(b-a)(b-c)} + \frac{c-a}{b-a} \quad (\text{A.37})$$

pudiendo comprobarse también que $F_{\tilde{x}_2}(b) = 1$. Igualando $F_{\tilde{x}_2}(x)$ a u , uniformemente distribuida ahora en $\left(\frac{c-a}{b-a}, 1\right)$ y despejando x se obtiene

$$x = b \pm \sqrt{(b-c)^2 - (b-c)[u(b-a) - (c-a)]} \quad (\text{A.38})$$

En resumen, se ha de generar una variable u uniformemente distribuida en $(0,1)$ y calcular x como

$$x = \begin{cases} a + \sqrt{u(c-a)(b-a)} & 0 < u \leq \frac{c-a}{b-a} \\ b - \sqrt{(b-c)^2 - (b-c)[u(b-a) - (c-a)]} & \frac{c-a}{b-a} < u < 1 \end{cases} \quad (\text{A.39})$$

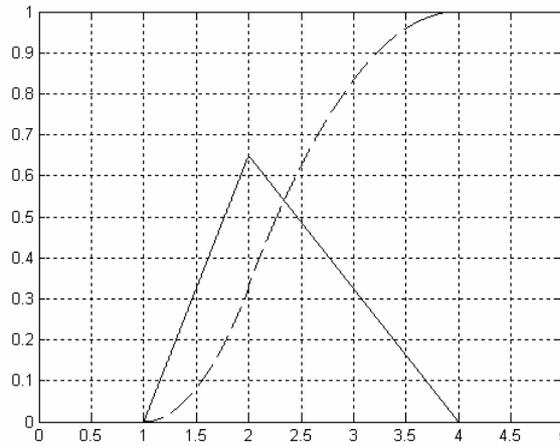


Figura A.5 – FDP y fdp triangulares de $a = 1$, $c = 2$ y $b = 4$.

A.7. Distribución de Rayleigh

La función de densidad de probabilidad de Rayleigh se define como

$$f_{\tilde{r}}(r) = \begin{cases} \frac{r}{\sigma^2} e^{-r^2/2\sigma^2} & r \geq 0 \\ 0 & r < 0 \end{cases} \quad (\text{A.40})$$

Puede demostrarse que

$$E\{r\} = \sqrt{\pi/2} \sigma \quad (\text{A.41})$$

$$E\{r^2\} = 2\sigma^2 \quad (\text{A.42})$$

La función de distribución se calcula como

$$F_{\tilde{r}}(r) = \Pr\{\tilde{r} \leq r\} = \int_0^r f_{\tilde{r}}(y) dy = \int_0^r \frac{y}{\sigma^2} e^{-y^2/2\sigma^2} dy = -e^{-y^2/2\sigma^2} \Big|_0^r = 1 - e^{-r^2/2\sigma^2} \quad (\text{A.43})$$

Igualando la función de distribución obtenida a una variable u uniforme entre 0 y 1

$$F_{\tilde{r}}(r) = 1 - e^{-r^2/2\sigma^2} = 1 - u = u \quad (\text{A.44})$$

$$e^{-r^2/2\sigma^2} = u \quad (\text{A.45})$$

$$r = \sqrt{-2\sigma^2 \ln u} \quad (\text{A.46})$$

con u también distribuida uniformemente en (0,1). Para evitar el signo ‘-’ dentro de la raíz cuadrada, también puede escribirse

$$r = \sqrt{2\sigma^2 \ln \frac{1}{u}} \quad (\text{A.47})$$

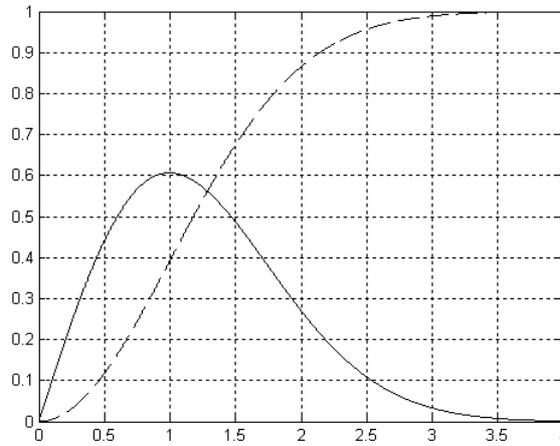


Figura A.6 – FDP y fdp de Rayleigh con $\sigma = 1$.

A.8. Distribución Normal

La generación de variables aleatorias de distribución normal o gaussiana es conceptualmente distinta a las estudiadas hasta este momento, que se basan en el método de la transformada inversa. La fdp normal es

$$f_{\tilde{x}}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-m_x)^2/2\sigma^2} \quad (\text{A.48})$$

Así, por definición, la función de distribución es

$$F_{\tilde{x}}(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-(t-m_x)^2/2\sigma^2} dt \quad (\text{A.49})$$

que no posee solución analítica, por lo que no puede aplicarse el método hasta ahora usado.

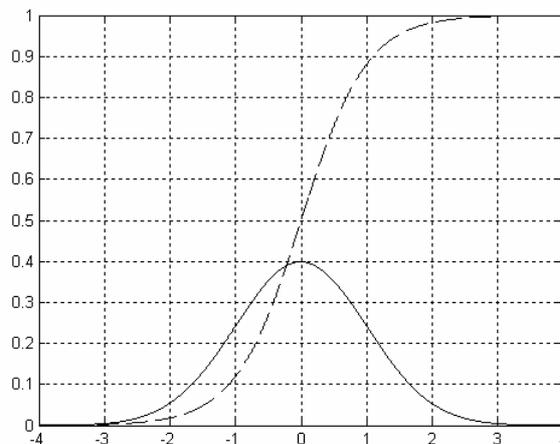


Figura A.7 – FDP y fdp normales o gaussianas de $m = 0$ y $\sigma = 1$.

Para calcularla se definen \tilde{x} e \tilde{y} dos variables aleatorias normales independientes de medias m_x y m_y respectivamente, y varianza σ . La función de densidad de probabilidad conjunta es separable

$$f_{xy}(x, y) = f_{\tilde{x}}(x) \cdot f_{\tilde{y}}(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-m_x)^2/2\sigma^2} \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-m_y)^2/2\sigma^2} = \frac{1}{\sigma^2 2\pi} e^{-\frac{(x-m_x)^2+(y-m_y)^2}{2\sigma^2}} \quad (\text{A.50})$$

Se realizará un cambio a coordenadas polares

$$\vec{r} = (x - m_x) + j(y - m_y) \quad (\text{A.51})$$

$$x = m_x + r \cos \theta \quad (\text{A.52a})$$

$$y = m_y + r \sin \theta \quad (\text{A.52b})$$

$$r = \sqrt{(x - m_x)^2 + (y - m_y)^2} \quad (\text{A.53a})$$

$$\theta = \arctan \frac{y - m_y}{x - m_x} \quad (\text{A.53b})$$

con $r \in [0, \infty)$ y $\theta \in [0, 2\pi)$. La relación entre distribuciones en ambos sistemas de coordenadas es

$$f_{r\theta}(r, \theta) = f_{xy}(x, y) \cdot \frac{\partial(x, y)}{\partial(r, \theta)} \quad (\text{A.54})$$

donde el jacobiano se calcula como

$$\frac{\partial(x, y)}{\partial(r, \theta)} = \begin{vmatrix} \frac{dx}{dr} & \frac{dy}{dr} \\ \frac{dx}{d\theta} & \frac{dy}{d\theta} \end{vmatrix} = \begin{vmatrix} \cos \theta & \sin \theta \\ -r \sin \theta & r \cos \theta \end{vmatrix} = r \quad (\text{A.55})$$

Tras realizar el cambio se tiene

$$f_{r\theta}(r, \theta) = \frac{r}{\sigma^2 2\pi} e^{-r^2/2\sigma^2} \quad (\text{A.56})$$

Integrando en r y θ se obtienen las funciones de densidad de probabilidad marginales:

$$f_{\tilde{r}}(r) = \int_0^{2\pi} f_{r\theta}(r, \theta) d\theta = 2\pi \cdot f_{r\theta}(r, \theta) = \frac{r}{\sigma^2} e^{-r^2/2\sigma^2} \quad (\text{A.57})$$

$$f_{\tilde{\theta}}(\theta) = \int_0^{\infty} f_{r\theta}(r, \theta) dr = \frac{1}{2\pi} \int_0^{\infty} \frac{r}{\sigma^2} e^{-r^2/2\sigma^2} dr = \frac{1}{2\pi} \left[-e^{-r^2/2\sigma^2} \right]_0^{\infty} = \frac{1}{2\pi} \quad (\text{A.58})$$

llegando a una distribución de Rayleigh para r y una uniforme entre 0 y 2π para θ . La generación de variables aleatorias con estas distribuciones ya ha sido explicada, por lo que las variables aleatorias normales se generan como

$$x = m_x \pm \sqrt{2\sigma^2 \ln \frac{1}{u}} \cdot \cos(2\pi \cdot v) \quad (\text{A.59a})$$

$$y = m_y \pm \sqrt{2\sigma^2 \ln \frac{1}{u}} \cdot \sin(2\pi \cdot v) \quad (\text{A.59b})$$

con u y v uniformemente distribuidas entre 0 y 1.

Acrónimos

- ALBA: *Agregated Least Busy Alternative*
- ATM: *Asynchronous Transfer Mode*
- BGP: *Border Gateway Protocol*
- CR-LDP: *Constraint-based Routed LDP*
- DAR: *Dynamic Alternative Routing*
- DCR: *Dynamically Controlled Routing*
- DNHR: *Dynamic Non-Hierarchical Routing*
- DoS: *Depth-of-Search*
- EDR: *Event Dependent Routing*
- EFP: *Erlang Fixed Point*
- FDM: *Frequency Division Multiplexing*
- FDP: *Función de Distribución de Probabilidad*
- fdp: *función de densidad de probabilidad*
- FR: *Fixed Routing*
- GoS: *Grade-of-Service*
- IEEE: *International Electrical and Electronic Engineers*
- IGP: *Interior Gateway Protocol*
- IP: *Internet Protocol*
- ITU-T: *International Telecommunications Union*
- LAW: *Lastabhängige Automatische Wegesuche*
- LBA: *Least Busy Alternative*
- LLR: *Least Loaded Routing*
- LSP: *Label-Switched Path*
- MPLS: *Multi-Protocol Label Switching*
- OSPF: *Open Shortest Path First*
- PASTA: *Poisson Arrival See Time Averages*
- PNNI: *Private Network-Network Interface*
- RDSI: *Red Digital de Servicios Integrados*
- RSVP: *Resource Reservation Protocol*
- RTNR: *Real-Time Network Routing*
- SDR: *State Dependent Routing*
- SS7: *Signaling System number 7*
- STR: *State-and-Time dependent Routing*
- TDM: *Time Division Multiplexing*
- TDR: *Time Dependent Routing*
- VoIP: *Voice-over-IP*
- WIN: *Worldwide Intelligent Network routing*

Bibliografía y Referencias

- [1] J. Bellamy, "Digital Telephony", Ed. John Wiley & Sons, NY 1991, Second Edition.
- [2] J.E. Flood, "Telecommunications Switching, Traffic and Networks", Ed. Prentice Hall, 1995.
- [3] D. Bertsekas, R. Gallager, "Data Networks", Prentice Hall, 1992, Second Edition.
- [4] G.R. Ash, P. Chemouil, "20 years of Dynamic Routing in Circuit-Switched Networks: Looking backward to the Future", Computer Communications Magazine, 2004.
- [5] G.R. Ash, "Performance Evaluation of QoS-Routing Methods for IP-Based Multiservice Networks", Computer Communications Magazine, July 2003.
- [6] W.M. Wong, K.M. Chang, Tak-Shing Peter Yum, "Analysis of Rerouting in Circuit-Switched Networks", IEEE/ACM Transactions on Networking, vol. 8, no. 3, June 2000.
- [7] W.K. Felix Lor, K.Y. Michael Wong, "Decentralized Neural Dynamic Routing in Circuit-Switched Networks", Applications of Neural Networks in Telecommunications, May 1995.
- [8] R. Gibbens, S. Turner, "Call routing in telephone networks", May 1997, <http://pass.maths.org.uk/issue2/dar/index.html>.
- [9] G. Trangmoe, "A Comparative Study of Dynamic Routing in Circuit-Switched Networks", December 1995, <http://www.statslab.cam.ac.uk/~frank/DAR/index.html>.
- [10] RFC 2386, E. Crawley, F. Nair, B. Rajagopalan, H. Sandick, "A Framework for QoS-Based Routing in the Internet", August 1998.
- [11] D. Awduche, "MPLS and Traffic Engineering in IP Networks", IEEE Com. Magazine, Vol. 17, no. 12, 1999.
- [12] ATM Forum, "Private Network-Network Interface Specification Version 1.1 (PNNI 1.1)", April 2002.
- [13] RFC 2702, D. Awduche et al, "Requirements for Traffic Engineering over MPLS", IETF, September 1999.
- [14] J. Davidson, J. Peters, "Fundamentos de voz sobre IP", PEARSON EDUCACIÓN, S.A., Madrid 2001.
- [15] G.R. Ash et al, "Routing Evolution in Multiservice Integrated Voice/Data Network", Proc. Of Int'l Teletraffic Congress, ITC 16, Edinburgh, UK, 1999.
- [16] H. Ahmadi et al, "Dynamic Routing and Call Control in High-Speed Integrated Networks", Proc. Of Int'l Teletraffic Congress, ITC 13, Copenhagen, Denmark, 1991.

