

TESIS

Presentada el día 30 de setiembre de 2011 en la

Universidad de La República, UdelaR

para obtener el título de

MAGÍSTER EN INGENIERÍA
MATEMÁTICA

para

María Elisa BERTINAT TULIPANO

Instituto de Investigación : LPE - IMERL

Componentes universitarios :

UNIVERSIDAD DE LA REPÚBLICA

FACULTAD DE INGENIERÍA

Título de la tesis :

*Estudio de métodos de reducción de varianza para modelos de
confiabilidad con restricciones de diámetro.*

TESIS DE MAESTRÍA

Comité de examinadores:

Dr. Franco	ROBLEDO	Director de Tesis
Dr. Héctor	CANCELA	Director de Tesis
Dr. Gerardo	RUBINO	Presidente
Prof. Ing. Enrique	CABAÑA	
Dr. Eduardo	CANALE	

Agradecimientos

Quiero dedicar esta tesis a todas las personas que desde diferentes lugares estuvieron conmigo brindándome su apoyo y comprensión.

Agradezco a mis tutores, Dr. Héctor Cancela y Dr. Franco Robledo, por guiarme en esta tesis. Gracias por compartir su conocimiento y acompañarme en este proceso.

Agradezco a mis compañeros de trabajo, Ingeniero Pablo Romero y Licenciado Darío Padula, por todos los momentos compartidos, en donde el aprendizaje y el trabajo fueron de la mano con la amistad.

Gracias a mi amiga Fernanda por su apoyo y sinceridad, presente en esta etapa como en tantas otras.

Gracias a mi novio y a mis padres por la comprensión y la compañía que me brindan en cada uno de los desafíos que inicio.

Índice

Índice de Contenidos	1
I INTRODUCCIÓN	5
1. Introducción	11
1.1. Motivación	11
1.2. Objetivo	12
1.3. Organización del documento	12
II MARCO TEÓRICO	15
2. Confiabilidad en redes	17
2.1. Introducción	17
2.2. Áreas de aplicación	18
2.3. Modelo estocástico	20
2.4. Cálculo de la medida de confiabilidad	23
2.4.1. Procedimientos de reducción de grafos	25
2.4.2. Métodos de cálculo exacto	27
2.4.2.1. Enumeración del espacio de configuraciones	27
2.4.2.2. Principio de inclusión y exclusión y Método de Poincaré	28
2.4.2.3. Métodos de Factorización	29
2.4.3. Métodos aproximados	32
2.4.3.1. Método Monte Carlo	33
2.4.3.2. Formulación alternativa de Monte Carlo Crudo	35
2.4.3.3. Métodos de reducción de varianza	38
3. Confiabilidad diámetro acotada	59
3.1. Introducción	59
3.2. Modelo probabilístico	60
3.3. Métodos de cálculo exacto	63
3.3.1. Procedimientos de reducción de grafos	65
3.3.2. Método de Factorización	67

3.3.3. Método Recursivo	70
3.3.4. Enumeración de pathsets	73
3.4. RVR: versión basada en muestreo multinomial	74
3.5. Antecedentes	81
3.6. Problemas abiertos	82
III ANÁLISIS DE ROBUSTEZ	83
4. Métodos de cálculo: análisis de eficiencia y robustez	85
4.1. Introducción	85
4.2. Cobertura e intervalos de confianza	86
4.2.1. Función de cobertura: análisis empírico	88
4.3. Propiedades asintóticas	89
5. Diagnóstico de robustez de RVR	91
5.1. Resultados numéricos: estudio de casos	92
5.1.1. Implementación	92
5.1.2. Casos de estudio	94
5.1.3. Notación	95
5.1.4. Resultados Empíricos	96
5.1.4.1. Grafo completo K_n	97
5.1.4.2. Grafo Grilla	98
5.1.4.3. Grafo Arpanet	99
5.1.4.4. Grafo Dodecahedro	100
5.1.4.5. Grafo Red ANTEL	101
5.1.5. Comentarios generales	103
5.2. Diagnóstico empírico	104
5.2.1. Función de Cobertura	104
5.2.1.1. Casos de prueba	105
5.2.2. Resultados	106
5.3. Estudio asintótico del Error Relativo	109
5.4. Conclusiones	113
IV CONCLUSIONES	115
6. Conclusiones Generales	117
6.1. Trabajo futuro	118
Bibliografía	123
Lista de Figuras	124

Índice	3
V ANEXOS	129
7. Algoritmos auxiliares	131
7.1. Algoritmo de búsqueda en profundidad	131
7.2. Algoritmo de Búsqueda en Anchura	134
7.3. Algoritmo de Búsqueda en Anchura para identificar si un grafo es D-conexo . .	135
7.4. Algoritmo de Ford-Fulkerson	138
7.5. Camino más corto con restricciones de largo: Algoritmo Dijkstra adaptado . . .	141
7.6. Muestra aleatoria Multinomial	144
8. Trabajo en el cluster	147
9. Glosario	149

Parte I

INTRODUCCIÓN

Abstract

El desarrollo de la tecnología en el área de las comunicaciones ha derivado diversas líneas de investigación. Entre ellas, la confiabilidad de la red, definida como la probabilidad de que la red permanezca operativa ante fallas eventuales de sus componentes.

El cálculo exacto de la probabilidad de funcionamiento de una red vulnerable a fallas es un problema combinatorio perteneciente a la clase NP -difícil. Usualmente se recurre a métodos aproximados para su estimación.

El método de Reducción Recursiva de la Varianza (RVR) es una alternativa, que se caracteriza por su bajo costo computacional, particularmente en redes donde la ocurrencia de fallas es un evento raro (probabilidad de falla cercana a 0).

Esta tesis se basa en el método RVR aplicado al modelo de arista-confiabilidad diámetro acotada (se asume falla solamente en los enlaces). Este modelo estudia la probabilidad de que la red comunique un conjunto de nodos objetivo a través de caminos de largo acotado. Es el caso de la comunicación por videoconferencia.

Se presenta el estado del arte de la medida de arista-confiabilidad clásica (sin restricciones de diámetro) y su generalización, la arista-confiabilidad diámetro acotada.

Se desarrolla el análisis cuantitativo de la robustez del método RVR para estimar la arista-confiabilidad diámetro acotada en escenarios donde la ocurrencia de fallas es un evento raro. Junto con las redes clásicas se incluye el estudio de la red de transporte de la empresa estatal de telecomunicaciones ANTEL.

Los resultados indican que si bien el método RVR , en la versión implementada, tiene desempeño computacional ampliamente satisfactorio, no es robusto en redes altamente confiables.

Palabras clave: confiabilidad diámetro acotada, RVR , eventos raros, robustez.

Notación General

- $G = (V, E)$: grafo no dirigido donde V es el conjunto de vértices y E el conjunto de enlaces de G .
- $K \subseteq V$: conjunto de nodos especiales de G , denominado conjunto de nodos objetivo o terminales.
- m : cardinal del conjunto de aristas de G : $|E| = m$.
- n : cardinal del conjunto de nodos de G : $|V| = n$.
- k : cardinal del subconjunto de nodos: $|K| = k$.
- e_i : i -ésima arista de E , siendo $E = e_1, e_2, \dots, e_m$.
- v_i : i -ésimo nodo en V , siendo $V = v_1, v_2, \dots, v_n$.
- r_i : probabilidad de que el enlace $i \in E$ esté operativo. Se denomina confiabilidad elemental.
- \vec{r} : vector de confiabilidad elemental, tal que $\vec{r}(i) = r_i$.
- q_i : probabilidad de que el enlace i esté en falla: $q_i = 1 - r_i$. Se denomina anticonfiabilidad elemental.
- $G * e$: Operación sobre el grafo G que indica la contracción de la arista e en G .
- $G - e$: Operación sobre el grafo G , que indica la eliminación de la arista e en G .
- $X = (X_1, X_2, \dots, X_m)$: vector aleatorio en \mathbb{R}^m donde X_i es el i -ésimo elemento de X tal que si $X_i = 1$ entonces e_i opera, y si $X_i = 0$ la arista e_i falla. Se denomina vector de estado o configuración de G .
- Ω : conjunto de todos los vectores de estado posibles de G .
- $\Phi(X)$: una función de estructura $\Phi(\cdot) : \{0, 1\}^m \rightarrow \{0, 1\}$, $|E| = m$. Una función es coherente si cumple:
 - Si todos los componentes del sistema operan (fallan) el sistema opera (falla).

- Si el sistema está operativo (falla) y el estado de un componente cambia de falla a operativo (de operativo a falla) el sistema permanece operativo (falla).
- Todos los componentes son relevantes (su estado afecta el estado global de la red).

Formalmente, una función coherente cumple:

- $\Phi(\vec{0}) = 0, \Phi(\vec{1}) = 1,$
- Si $\vec{x} < \vec{y}$ entonces $\Phi(\vec{x}) \leq \Phi(\vec{y}),$
- Para cada componente i existe algún vector de estado \vec{x} tal que $\Phi(\vec{x}, 0_i) \neq \Phi(\vec{x}, 1_i),$

donde

- $\vec{0}$ es un vector de estados con todos sus entradas igual a 0 ,
 - $\vec{1}$ es un vector de estados con todos sus entradas igual a 1,
 - $\vec{x} \leq \vec{y}$ la relación $x_i \leq y_i, \forall i,$
 - $\vec{x} < \vec{y}$ la relación $\vec{x} \leq \vec{y}$ con algún $j: x_j < y_j,$
 - $(\vec{x}, 0_i)$ es el vector de estado que se construye a partir de \vec{x} fijando x_i en 0 y
 - $(\vec{x}, 1_i)$ es el vector de estado que se construye a partir de \vec{x} fijando x_i en 1.
- $1_{\{E\}}$: Función indicatriz del evento E tal que:

$$1_{\{E\}} = \begin{cases} 1, & \text{si } E \text{ es verdadero,} \\ 0, & \text{en el otro caso} \end{cases}$$

- $R_K(G)$: Medida de confiabilidad K –terminal del grafo G con conjunto de nodos objetivo K . Asumiendo que los nodos de G son perfecto (no fallan) y que los enlaces fallan de manera independiente con probabilidad dada, la medida $R_K(G)$ es la probabilidad de que los nodos de K permanezcan comunicados una vez eliminados los enlaces en falla.
- D : es la cota superior del largo de los caminos operativos que comunican los nodos objetivo de G .
- $R_{K,D}(G)$: medida de confiabilidad diámetro acotada del grafo G con conjunto de nodos objetivo K y cota D . Asumiendo nodos perfectos, es la probabilidad de que los k nodos objetivo permanezcan comunicados con caminos de largo D o menor.
- $Q_K(G) = 1 - R_K(G)$: medida de anticonfiabilidad del grafo G con conjunto de nodos objetivo K , se define como la probabilidad de que la red no opere.
- $Q_{K,D}(G) = 1 - R_{K,D}(G)$: medida de anticonfiabilidad diámetro acotada del grafo G con conjunto de nodos objetivo K y cota D .

Capítulo 1

Introducción

1.1. Motivación

El desarrollo de las redes de comunicación, la diversificación de los servicios y funcionalidades que ofrecen y el número creciente de usuarios, han impulsado distintas líneas de investigación dedicadas al análisis y diseño de redes.

La confiabilidad en redes, temática en la que se centra este trabajo, es una de las líneas de investigación cuyo objeto de estudio es la capacidad que tiene la red de mantenerse operativa ante fallas en sus componentes.

La red se representa como un conjunto de nodos, que indican los puntos que se desean comunicar, y un conjunto de enlaces, que indican las líneas a través de las cuales se conectan dichos nodos. Las fallas en los componentes, resultado de daños climáticos, daños intencionales causados por el hombre, entre otros, evidencian la vulnerabilidad de la red, que eventualmente falla ante la no operatividad de alguno de sus enlaces o nodos.

Formalmente, los componentes de la red tienen probabilidad positiva de falla y la medida de confiabilidad coincide con la probabilidad de que la red esté operativa. En este trabajo se asume un modelo probabilístico de arista confiabilidad, donde los nodos son perfectos y los enlaces fallan con probabilidades conocidas e independientemente los unos de los otros.

Generalmente el concepto de comunicación exitosa supone la posibilidad de contactar dos o más puntos en la red. No obstante existen situaciones donde se requiere además que los lapsos de tiempo durante el cual ocurre la comunicación o el largo de los caminos a través de los que se realiza no superen una tolerancia dada. Un ejemplo es la comunicación por video conferencia donde es fundamental que los mensajes de voz e imagen transferidos no sufran retardos importantes que entorpezcan el diálogo en tiempo real. En este contexto, se define una nueva medida denominada confiabilidad diámetro acotada que incorpora una cota superior al largo de los caminos a través de los cuales se desarrolla la comunicación. La red se considera en estado operativo si los nodos se conectan por caminos de largo no mayor a una cota D , una vez eliminados los componentes en falla.

La medida de confiabilidad diámetro acotada generaliza el concepto de confiabilidad clásica adaptándose a un número mayor de problemas en donde existen restricciones de tiempo o

distancia en la comunicación. Puede considerarse como un indicador de la calidad de comunicación así como un parámetro de comparación, que en conjunto con otras dimensiones, brinda una pauta para seleccionar un diseño de red sobre otro.

Conceptualmente la medida de confiabilidad es sencilla, sin embargo el problema de cálculo es complejo, perteneciente a la familia de problemas *NP*-difíciles.

1.2. Objetivo

El objetivo de este trabajo es analizar el desempeño de un método de Reducción Recursiva de la Varianza (*RVR*) para estimar la medida de confiabilidad diámetro acotada. Este método recursivo presenta ventajas respecto del método Monte Carlo Crudo (*CMC*) en distintos casos de análisis, particularmente cuando se aplica a redes altamente confiables. El objetivo de esta tesis es:

- Comparar el desempeño del método *RVR* con el método *CMC*, tomando como parámetros de referencia la precisión en las estimaciones, cuantificada en términos de la varianza de la estimación y el tiempo de cálculo.
- Estudiar la robustez del método *RVR* respecto de la magnitud de la probabilidad de falla de los enlaces de la red. En otras palabras, el objetivo es determinar si la calidad de las estimaciones obtenidas con *RVR* se deteriora cuando la falla de los enlaces se hace menos probable.

1.3. Organización del documento

En el Capítulo 1 se presenta una introducción general al problema de confiabilidad en redes y el objetivo de esta tesis.

En el Capítulo 2 se desarrolla el concepto clásico de Confiabilidad en Redes. En particular se detallan áreas de aplicación de la medida de confiabilidad. Se desarrolla el modelo probabilístico que representa el comportamiento de falla de la red y la complejidad asociada al cálculo de la medida. Por último se enumeran algoritmos de cálculo exacto y aproximado para determinar la medida de confiabilidad.

En el Capítulo 3 se presenta la medida de confiabilidad en redes diámetro acotada, una generalización de la medida de confiabilidad clásica presentada en el Capítulo 2. Se desarrolla el modelo probabilístico, obtenido como una extensión del modelo clásico para el nuevo concepto de operatividad de la red bajo la restricción de diámetro. Para finalizar se consideran métodos de cálculo de la medida, con especial énfasis en una versión multinomial del método Recursivo de Reducción de Varianza *RVR*.

En el Capítulo 4 se especifican propiedades teóricas asociadas a la robustez de los estimadores en contexto de eventos raros basadas en el error relativo de las estimaciones y la cobertura de los intervalos de confianza empíricos. En contexto de eventos raros es deseable que el error

relativo permanezca acotado y que la cobertura empírica de los intervalos sea cercana a la cobertura teórica.

En el Capítulo 5 se desarrolla un análisis cuantitativo del método RVR para el cálculo de la confiabilidad diámetro acotada. Se contrasta este método con el método Monte Carlo Crudo en función de estimaciones obtenidas para casos de prueba. Se detallan resultados empíricos que analizan la robustez de la versión de RVR implementada en función de la cobertura de intervalos de confianza. Dichos resultados se complementan con el estudio asintótico del error relativo para un caso puntual.

En el Capítulo 6 se presentan las principales contribuciones de esta tesis.

El Capítulo 7 es un anexo que detalla un conjunto de algoritmos auxiliares aplicados en esta tesis. En el Capítulo 8 se detalla el Script utilizado en todas las simulaciones implementadas en el cluster.

Por último el Capítulo 9 presenta un Glosario de conceptos básicos.

Parte II

MARCO TEÓRICO

Capítulo 2

Confiabilidad en redes

2.1. Introducción

La medida de confiabilidad es una medida de seguridad de funcionamiento con amplia aplicación en la actualidad, utilizada para evaluar redes eléctricas, sistemas de transporte urbano, redes de interconexión y sistemas de comunicación, siendo esta última una de las aplicaciones centrales [1]. La medida permite cuantificar la calidad de comunicación de la red en función de su capacidad de mantenerse operativa en sucesos puntuales como fallas originadas por catástrofes climáticas.

Un ejemplo asociado al diseño de redes, en donde la confiabilidad es un aspecto de relevancia, es el proyecto de diseño de red de la empresa estatal telefónica ANTEL (Administración Nacional de Telecomunicaciones) cuyo fin es aumentar la velocidad de acceso a Internet incrementando el ancho de banda. De acuerdo con lo informado por la prensa uruguaya en julio del año 2009, el proyecto de construcción de un cable submarino de ANTEL en sociedad con Telecom Argentina, implica el tendido submarino de fibra óptica entre ciudades de ambos países, lo que ofrece ventajas respecto a opciones terrestres en lo que refiere a seguridad y confiabilidad ante eventuales fallas.

Para representar la red se considera una estructura de grafo estocástico con dos estados posibles: operativo o falla. Su estado se describe mediante un modelo probabilístico que asigna una probabilidad de falla a cada uno de sus componentes.

En el caso de una red de comunicación los nodos representan los computadores de los usuarios, servidores, routers, etc., los enlaces representan las líneas de comunicación entre nodos y un enlace operativo equivale a la posibilidad de que dos o más usuarios puedan completar la transmisión de información.

El problema de evaluar la confiabilidad ha recibido gran atención por la comunidad de investigadores, tanto por la importancia de brindar servicios de calidad como por la complejidad computacional de su cálculo [2–4]. Perteneciente a la clase de problemas de complejidad *NP*-difíciles [5, 6], el problema de cálculo exacto de la medida de confiabilidad en redes es

computacionalmente prohibitivo [1], motivando un continuo esfuerzo por encontrar métodos de resolución más eficientes.

En este capítulo se presenta el estado del arte de la medida de confiabilidad clásica. En la Sección 2.2 se presenta una reseña de áreas de aplicación de la medida de confiabilidad. En la Sección 2.3 se define la medida de confiabilidad en redes y se presenta el modelo probabilístico asociado a la red.

En la Sección 2.4 se desarrolla la complejidad en torno al cálculo de la medida de confiabilidad. En particular se detallan métodos de cálculo exacto y métodos aproximados. Dentro de los métodos aproximados se realiza especial énfasis en métodos de reducción de varianza, aplicación en la que se centra esta tesis.

2.2. Áreas de aplicación

Los modelos basados en redes permiten representar en forma sencilla una variedad de problemas en contextos diversos como redes de comunicación de voz y/o datos, redes de transporte, redes de energía eléctrica y sistemas de comando y control.

En lo que respecta a la medida de confiabilidad, muchos de los trabajos iniciales se desarrollan en contextos de alta confiabilidad, enfocados en situaciones donde una falla del sistema puede causar daños masivos o pérdidas de vidas humanas. Es el caso de sistemas de aviación o sistemas de control de reactores nucleares [1].

Además de la importancia en términos sociales, los sistemas confiables son de relevancia en términos financieros como sistemas bancarios, sistemas de verificación de créditos, redes de comunicación.

En trabajos más recientes el foco de la investigación se centra en el diseño de redes sujetas a estándares de confiabilidad. La validación de un diseño de red, entre otros aspectos, depende de la capacidad de mantenerse operativa prefiriendo diseños que optimicen la confiabilidad [7–9].

A continuación se presentan algunas áreas de aplicación. (Ver [1] por un desarrollo más detallado)

1. Redes de Conmutación de Paquetes a Nivel de Backbone

Este tipo de redes fueron desarrolladas en los años 1960 para permitir inter-conectividad entre numerosos usuarios a alta velocidad. ARPANET fue la primera red importante de este tipo que inspiró además nuevos trabajos en los comienzos de los años 1970.

El protocolo de ruteo que aplican las redes de conmutación de paquetes es dinámico, lo que permite utilizar caminos alternativos ante fallas en sus componentes. En este tipo de redes la medida de confiabilidad adoptada es una medida de conectividad. Sin embargo aplicar enrutamiento de tráfico sustituyendo enlaces fallados por nuevos enlaces puede generar congestión en las conexiones perjudicando el desempeño global de la red.

Con el avance de la tecnología y con el crecimiento del número de usuarios, las sucesoras de ARPANET son redes de mayor densidad, lo que asegura una menor probabilidad de falla del sistema. Además de la capacidad de conexión, se incorporan parámetros de calidad como la capacidad de los enlaces y el congestionamiento de la red, obteniendo una medida de desempeño global.

2. **Redes de Conmutación de Circuitos a Nivel de Backbone**

La existencia de grandes redes de telecomunicación son posibles gracias a las redes de conmutación de circuitos. Son las primeras redes de conmutación inventadas y se usan desde el surgimiento de la central telefónica automática, que data de 1879 [10]. En éste tipo de redes un canal se utiliza para conectar un par de usuarios en forma directa durante el tiempo que dura la llamada. Una falla en un componente de la red reduce los canales de comunicación disponibles y en consecuencia los circuitos de llamada para los usuarios. Es lo que se conoce como llamada bloqueada. Los trabajos más recientes en redes de conmutación de circuitos asumen que un enlace falla si está bloqueado y miden el desempeño en términos del bloqueo.

3. **Redes de Interconexión**

Las redes de interconexión son un caso especial de las redes de conmutación (punto 2), aplicadas para conectar procesadores en paralelo y memorias de computadores de arquitectura paralela. La existencia de máquinas paralelas con diversos procesadores plantea la necesidad de comunicación y las operaciones de intercambio de información son realizadas gracias a una red de interconexión

Los modelos de conectividad usados consideran la falla de los componentes y su congestión.

4. **Redes de Fibra óptica**

En 1977, luego de años de investigación, se envió la primera transmisión telefónica a través de fibra óptica. Éstas redes transmiten la comunicación por canales mediante ondas de luz a través de fibra de vidrio, un material nuevo que permitió incrementar la capacidad de transmisión respecto a las redes tradicionales y construir redes más densas. Éstas nuevas cualidades repercutieron positivamente en los niveles de confiabilidad.

5. **Arquitectura de conmutación y cómputo tolerante a fallas**

Los sistemas de cómputo tolerantes a fallas son computadores que continúan operando en presencia de falla de uno o más de sus componentes. Fueron usados en los años 1970 como la base de los equipos de conmutación en telecomunicaciones. Más recientemente se han incorporado a computadores de arquitectura paralela.

Éste tipo de arquitectura consta de múltiples componentes del mismo tipo para incrementar el rendimiento del sistema. En general la tolerancia a falla y el sistema de cómputo paralelo son modelados como redes para analizar su confiabilidad, que se asocia a medidas de capacidad.

6. **Propagación aleatoria del fuego**

En este contexto se emplea un modelo en donde la probabilidad de falla se interpreta

como la probabilidad de que el fuego pueda propagarse a través de las barreras (paredes por ejemplo). Una vez que el fuego se establece en una habitación o un edificio existe la posibilidad de que el fuego se propague a través de barreras a habitaciones o edificios adyacentes [11].

7. Redes de Transporte

Las redes de transporte son altamente vulnerables a fallas como la rotura de un puente o el cierre de calles. En casos extremos, como desastres climáticos pueden generarse interrupciones prolongadas afectando el tráfico diario. En este sentido mejorar la confiabilidad de las redes de transporte es primordial para minimizar los daños generados por desastres.

La confiabilidad en redes de transporte puede ser modelada considerando exclusivamente la topología de la red, o bien, la topología y el flujo en los enlaces. En el segundo caso además de conectividad deben mantenerse determinados indicadores como niveles de flujo, tiempo de viaje, capacidad, entre otros [12, 13].

2.3. Modelo estocástico

La confiabilidad se define como la probabilidad de que una red se encuentre en estado operativo una vez eliminados los componentes en falla y puede interpretarse como una medida de calidad de funcionamiento del servicio que proporciona.

La red se representa topológicamente mediante un grafo no dirigido probabilístico sin lazos $G = (V, E)$, con V y E el conjunto de nodos y enlaces respectivamente. La medida de confiabilidad asociada depende del conjunto de nodos $K \subseteq V$ que se desea comunicar, del modelo probabilístico que describe su funcionamiento y la probabilidad de falla de los componentes.

En función del conjunto de nodos K que se requiere comunicar se distinguen tres casos de aplicación de la medida de confiabilidad del grafo G .

Definición 2.3.1 Dado un grafo $G = (V, E)$, el conjunto de nodos objetivo $K \subseteq V$, y $2 \leq |K| \leq n = |V|$, nodos perfectos y enlaces que fallan con probabilidad dada e independientemente unos de los otros, la **medida de confiabilidad K -terminal**, $R_K(G)$, es la probabilidad de que todo par de nodos en K esté conectado por al menos un camino operativo de G .

Se distinguen dos casos especiales de la medida $R_K(G)$: $K = \{s, t\}$ y $K = V$. En particular el caso $K = V$ se conoce como *broadcasting* en redes de comunicación. En ambos casos el cálculo de la medida es un problema de complejidad *NP*-difícil [1, 3].

Definición 2.3.2 Dado $G = (V, E)$, $\{s, t\} \in V$ la **medida de confiabilidad s - t -terminal**, $R_{\{s,t\}}(G)$, es la probabilidad de que exista al menos un camino operativo entre los nodos s y t .

Definición 2.3.3 Dado $G = (V, E)$, la **medida all-terminal o V -terminal**, $R_V(G)$, es la probabilidad de que exista al menos un camino operativo entre todo par de nodos de G .

Se define además el complemento de la medida de confiabilidad, la probabilidad de falla de la red: $Q_K(G) = 1 - R_K(G)$, denominada **medida de anticonfiabilidad de G** .

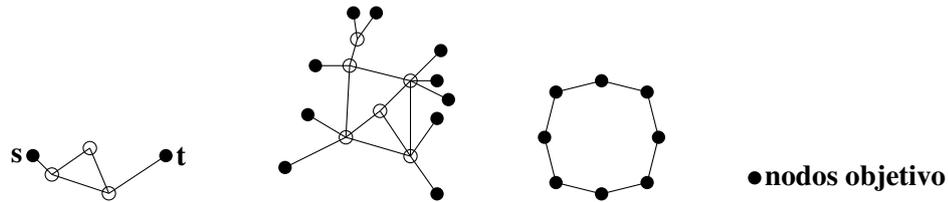


Figura 2.1: Confiabilidad s - t terminal, K - terminal y V - terminal

De acuerdo con el comportamiento de falla asumido en los componentes de la red se distinguen tres modelos probabilísticos: modelo de *nodo confiabilidad*, en el que se asume falla únicamente en los nodos, de *arista confiabilidad*, donde se asume falla sólo en los enlaces, ó de *nodo-arista confiabilidad* donde se asume falla en ambos componentes. Esta tesis se centra en el modelo arista confiabilidad o de confiabilidad clásico, uno de los más utilizados en la literatura [1].

Para cualquier conjunto de nodos objetivo cada enlace de G se describe como una variable aleatoria Bernoulli cuyos estados posibles son operativo o falla. La probabilidad de funcionamiento de la red depende del estado de cada uno de sus componentes. Implícitamente el modelo admite la utilización de caminos alternativos para mantener la red operativa ante la falla de algún enlace.

Asociada a la variable aleatoria Bernoulli recién mencionada se define la probabilidad de operatividad de la arista $e \in E$ o **confiabilidad elemental** de la arista e , r_e , y su complemento q_e , la probabilidad de falla o **anticonfiabilidad elemental**.

Cada configuración posible de los $|E| = m$ componentes del sistema es representado por un vector de estado aleatorio $X = (X_1, X_2, \dots, X_m)$, con $X \in \{0, 1\}^m$ y X_i el estado del i -ésimo enlace en E y $X_i = 1_{\{i \in E \text{ opera}\}}$. Se cumple:

$$\Pr(X_i = 1) = r_i.$$

Nótese que los vectores de estado de la red no son independientes entre sí dado que pueden compartir enlaces.

Asociado a X , se define una función binaria $\Phi(X)$ denominada función de estructura:

$$\Phi(\cdot) : \{0, 1\}^m \rightarrow \{0, 1\},$$

que asocia a cada X el valor 0 o 1, indicando que la red falla o está operativa de acuerdo con la configuración X :

$$\Phi(X) = 1_{\{\text{dado } X, G \text{ opera}\}}. \quad (2.1)$$

La medida de confiabilidad K -terminal y su complemento se definen como

$$R_K(G) = \Pr(\Phi(X) = 1), \quad Q_K(G) = \Pr(\Phi(X) = 0)$$

respectivamente. Se puede demostrar que la función Φ es coherente.

Cada vector de estado o configuración de G se clasifica como pathset o cutset de acuerdo con el valor que toma la función Φ evaluada en X .

Definición 2.3.4 Un *pathset* P es un conjunto de enlaces tal que si todos sus componentes operan entonces la red opera. Si x es un vector de estado tal que $x_j = 1, \forall x_j \in P$ entonces $\Phi(x) = 1$.

Definición 2.3.5 Un *minpath* P es un pathset minimal, en el sentido que la eliminación de cualquier elemento del minpath genera un sistema en falla. Si x es tal que $x_j = 1, \forall x_j \in P$ y $x_i = 0, \forall x_i \notin P$, entonces $\Phi(x) = 1$ y $\Phi((x, 0_j)) = 0$.

Definición 2.3.6 Un *cutset* C es un conjunto de enlaces cuya eliminación asegura la falla de la red. Si x es un vector de estado tal que $x_j = 0, \forall x_j \in C$ entonces $\Phi(x) = 0$.

Definición 2.3.7 Un *mincut* C es un cutset minimal, en el sentido que si al menos un elemento de C no se elimina de la red ésta permanece operativa. Si x es un vector de estado tal que $x_j = 0, \forall x_j \in C$ y $x_k = 1, \forall x_k \notin C$, entonces $\Phi(x) = 0$ y $\Phi((x, 1_j)) = 1$.

En el modelo de confiabilidad s - t -terminal un *pathset* es un camino elemental (no repite nodos intermedios) entre s y t y en el modelo K -terminal es un K -árbol recubridor de G .

Sea $P = \{P_1, P_2, \dots, P_{|P|}\}$ el conjunto de todos los *minpaths* de G y $\{E_i\}$ el evento “todos los enlaces de P_i operan”; puede definirse entonces la confiabilidad de la red en términos de E_i . La probabilidad de que la red opere es la probabilidad de que ocurra al menos uno de los eventos $\{E_i\}$:

$$\Pr(\Phi(X) = 1) = R_K(G) = \Pr(E_1 \circ E_2 \circ \dots \circ E_{|P|}), \quad (2.2)$$

Sea $C = \{C_1, C_2, \dots, C_{|C|}\}$ el conjunto de los *mincuts* de G y $\{F_i\}$ el evento “todos los enlaces de C_i fallan”. La probabilidad de que la red falle es la probabilidad de que ocurra al menos uno de los eventos $\{F_i\}$.

$$\Pr(\Phi(X) = 0) = Q_K(G) = \Pr(F_1 \circ F_2 \circ \dots \circ F_{|C|}).$$

Para finalizar esta sección se presenta un ejemplo sencillo considerando el grafo puente G de la Figura 2.2.

Ejemplo 2.1 Considere el grafo puente $G = (V, E)$ con $V = \{s, u, v, t\}$, $E = \{tv, tu, sv, su, vu\}$ y nodos objetivo $K = \{s, t\}$.

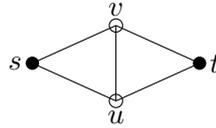


Figura 2.2: Grafo puente

Cada $e \in E$ tiene asignada una probabilidad de operar, $r_e \forall e \in E$ y $K = \{s, t\}$.
 Los pathsets de G son:

$$P = \{(s, u, t), (s, v, t), (s, u, v, t), (s, v, u, t)\}.$$

Si al menos uno de los pathsets en P opera entonces la red opera.

El vector de estado $X = (X_{su}, X_{sv}, X_{vu}, X_{tv}, X_{tu})$ asociado a G pertenece al conjunto de 2^5 configuraciones, enumeradas a continuación. Cada fila de las matrices representadas corresponde a una configuración de la red. El primer elemento de la matriz de la izquierda $(0, 0, 0, 0, 0)$ es el vector de estado con todas sus entradas igual a 0.

Configuraciones de falla de G	Configuraciones operativos de G
$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \end{bmatrix}$

2.4. Cálculo de la medida de confiabilidad

Las posibilidades de cálculo existentes para determinar el valor de la confiabilidad son diversas y dependen de las características del problema de interés. Uno de los aspectos principales que condicionan la forma de cálculo de la medida de confiabilidad $R_K(G)$ es el tamaño de la red: si es pequeña es posible aplicar métodos de cálculo exacto mientras que para redes

grandes deben considerarse, en general, métodos aproximados. El incremento del tamaño de la red influye en el tiempo computacional de los métodos exactos, haciéndolo crecer en forma exponencial en el número de nodos independientemente del valor de r_e [14]. En particular el problema de calcular la confiabilidad pertenece a la clase de complejidad NP -difícil, incluso en redes con idéntica probabilidad de operación en las aristas, y salvo estructuras puntuales, el cálculo exacto es computacionalmente prohibitivo [1].

Ball et al [1] presentan una noción que permite entender la complejidad del problema de confiabilidad en redes. El cálculo de la confiabilidad es un problema que tiene asociado la identificación y conteo de todas las configuraciones de la red que dan lugar a un estado operativo de G . El análisis de la *confiabilidad* depende de la estructura de la red y de la probabilidad de falla de sus componentes mientras que el análisis de la *complejidad* involucra conceptos utilizados en problemas de reconocimiento y optimización.

Para relacionar el análisis de complejidad a modelos de confiabilidad se considera el caso en que la confiabilidad elemental de los componentes es única, $r_e = r \forall e \in E$, y se escribe $R_K(G)$ como el *polinomio de confiabilidad* siguiente:

$$R_K(G) = \sum_{i=0}^M F_i r^{m-i} (1-r)^i \quad (2.3)$$

donde el factor F_i es la cantidad de conjuntos de enlaces S de cardinal i , tal que si S falla entonces el sistema opera, formalmente $F_i = |\{S : |S| = i; \Phi(x) = 1, \forall x \in \{E - S\}\}|$. Cada término del polinomio es la probabilidad de que exactamente $m - i$ componentes operen y den lugar a un estado operativo del sistema.

El problema asociado a la Ecuación 2.3, denotado como problema de *análisis funcional de la confiabilidad*, implica determinar cada factor F_i y es un problema de reconocimiento y conteo de configuraciones de la red (en [15, 16] se estudia el problema de determinar F_i).

Cada problema de conteo es al menos tan difícil como el correspondiente problema de reconocimiento. En algunos casos el problema de reconocimiento es de tiempo polinomial pero el problema de conteo correspondiente es $\#P$ -Completo. En efecto, en teoría de la complejidad computacional, la clase de complejidad $\#P$ es el conjunto de los problemas de conteo asociados a los problemas de decisión en el conjunto NP [1].

Sin embargo en estructuras triviales puede aplicarse resolución exacta. En redes pequeñas la enumeración de todas las configuraciones posibles de la red es relativamente sencilla; en redes donde algún nodo objetivo es aislado $R_K(G) = 0$ sin importar el tamaño de la red.

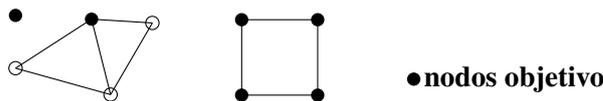


Figura 2.3: Ejemplo: nodo objetivo aislado y grafo trivial C_4

Para otro tipo de topologías se conocen algoritmos eficientes de cálculo exacto cuando las probabilidades de operación de las aristas son idénticas. Es el caso de los grafos completos,

grafos bipartitos, ciclos entre otros.

En el caso del grafo completo con n nodos K_n y la medida V -terminal asociada sea $A_n = R_V(K_n)$. Suponiendo que la confiabilidad elemental es la misma para todos los enlaces, q , se tiene [3]:

$$A_n = 1 - \sum_{j=1}^{n-1} C_{j-1}^{n-1} A_j q^{j(n-j)}, \quad (2.4)$$

donde A_1 coincide con la medida de confiabilidad asociada a un grafo K_1 que consta de un único nodo y por lo tanto $A_1 = 1$.

Para el caso s - t -terminal en el grafo completo K_n se define $T_n = R_{\{s,t\}}(K_n)$. En este caso la relación para el cálculo de la medida verifica [3]:

$$T_n = \sum_{j=2}^n C_{j-2}^{n-2} A_j q^{j(n-j)}. \quad (2.5)$$

Por último para el caso K -terminal de K_n , si se define $B_n = R_K(K_n)$ se cumple [3]:

$$B_n = \sum_{j=k}^n C_{j-k}^{n-k} A_j q^{j(n-j)}. \quad (2.6)$$

En el caso V -terminal de un grafo completo bipartito $K_{n,m}$ con $V = I_n \cup I_m$, $|I_n| = r$, $|I_m| = t$ e I_n, I_m dos conjuntos de nodos independientes. Si se denota $A_{i,j} = R_V(K_{n,m})$ se verifica [3]:

$$R_V(K_{n,m}) = \sum_{i=r}^n \sum_{j=t}^m C_{i-r}^{n-r} C_{j-t}^{m-t} A_{i,j} q^{i(m-j)} q^{j(n-i)}, \quad (2.7)$$

donde r es el número de nodos objetivo en I_n y t el número de nodos objetivo en I_m .

2.4.1. Procedimientos de reducción de grafos

Los procedimientos de reducción de grafos son aplicados en métodos de cálculo exacto y aproximado de la medida de confiabilidad $R_K(G)$. Permiten reducir el número de operaciones mediante la simplificación de la topología del grafo en estudio. El objetivo de este tipo de procedimiento es eliminar redundancia presente en la topología de la red, identificando una estructura equivalente a la original en términos de la medida de confiabilidad. Algunos de los cambios sobre la topología de la red son eliminación o contracción de aristas, fusión de nodos y reducción de caminos.

A continuación se presentan los procedimientos mencionados y la relación existente entre la confiabilidad de la red original y de la red transformada.

Las distintas transformaciones tienen asociado un costo computacional de orden polinomial y son aplicables en la mayoría de las estructuras. Aquellos grafos que luego de sucesivas reducciones serie-paralelo se reducen a un K -árbol (Árbol recubridor de K , $K \subseteq V$) se denominan grafos $s - p$ reducibles, en caso contrario grafos complejos.

1. **Contracción de una arista:**

Dada una arista $e = (x, y) \in E$ de $G = (V, E)$, la contracción de e implica la fusión entre sus nodos extremos. El grafo resultante, $G * e$ cumple $G_e = (V', E')$ donde $V' = V - \{x\}$ y $E' = E - \{e\}$. Cada una de las aristas de la forma (z, x) se sustituye por (z, y) .

Cualquier enlace puente de un grafo G sin enlaces irrelevantes es un enlace relevante. Un enlace $e \in E$ es un puente si separa dos nodos de la misma componente, i.e. $G \setminus \setminus e$ es desconexo.

Dado un grafo $G = (V, E)$ con un puente $e \in E$ con confiabilidad elemental r_e , la contracción de la arista puente preserva la confiabilidad con un factor multiplicativo r_e : $R_K(G) = r_e R_K(G * e)$.

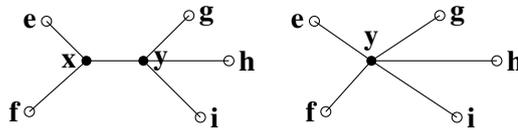


Figura 2.4: Contracción de una arista: G y $G * e$

2. **Eliminación de aristas irrelevantes:**

La eliminación de aristas en contexto de confiabilidad, se aplica a aquellas aristas irrelevantes que no afectan el valor de la medida. Dado un grafo $G = (V, E)$ y e una arista irrelevante, el grafo obtenido al eliminar e es $(G - e) = (V, E')$, donde $E' = E - \{e\}$. Dado que la arista es irrelevante se cumple $R_K(G) = R_K(G - e)$. Si se considera el grafo de la Figura 2.5 con $K = \{x, y\}$, todos los enlaces del conjunto $E \setminus (x, y)$ son irrelevantes.

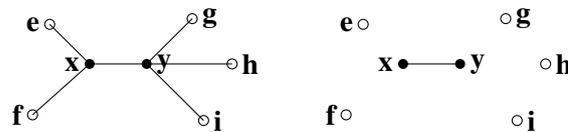


Figura 2.5: Eliminación de una arista: G y $G - e$

Aquellos enlaces que no pertenecen a ningún minpath del grafo se consideran irrelevantes, pues su estado no afecta el estado global de la red. Para cualquier medida de confiabilidad los lazos son siempre irrelevantes. En el caso $s-t$ y K terminal cualquier enlace que tenga un nodo extremo en un componente 2-conexo sin nodos terminales también es irrelevante [1].

3. Reducción en paralelo:

Dado $G = (V, E)$ y $\{e_1, e_2\} \in E$ dos aristas, se dice que ambas están en paralelo si $e_1 = (x, y)$ y $e_2 = (x, y)$. La reducción consiste en reemplazar las aristas que están en paralelo, e_1 y e_2 , por una arista $e = (x, y)$ preservando la confiabilidad de la red. En el nuevo grafo se cumple:

$$r_e = r_1 + r_2 - r_1 r_2.$$

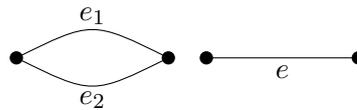


Figura 2.6: Reducción en paralelo

4. Reducción en serie:

Dado $G = (V, E)$ y $\{e_1, e_2\} \in E$ dos aristas, se dice que ambas están en serie si son incidentes a un mismo vértice de grado 2. La reducción en serie consiste en reemplazar las aristas en serie, $e_1 = (x, y)$ y $e_2 = (y, z)$, por una arista $e = (x, z)$, tal que la confiabilidad elemental del nuevo enlace cumple:

$$r_e = r_1 r_2.$$



Figura 2.7: Aristas en serie

2.4.2. Métodos de cálculo exacto

En esta sección se presentan métodos de cálculo exacto para obtener la medida de confiabilidad $R_K(G)$. Los distintos procedimientos se ilustran para el caso $K = \{s, t\}$, y la correspondiente medida $R_{\{s,t\}}(G)$, donde estados operativos de la red coinciden con arborescencias de raíz s .

Los distintos métodos tienen costo computacional exponencial en el tamaño de la red salvo algunas excepciones identificadas en clases particulares de grafos, donde el tiempo de cómputo es polinomial [1].

2.4.2.1. Enumeración del espacio de configuraciones

Uno de los métodos más simples para calcular la medida $R_K(G)$ es la enumeración del espacio de configuraciones de G , Ω_G . Su aplicación implica identificar todas las configuraciones posibles de la red y para cada uno evaluar si da lugar a un estado operativo o *pathset* y calcular

su probabilidad de ocurrencia.

Si bien la técnica conceptualmente es simple el costo computacional requerido para su implementación crece exponencialmente con el tamaño de la red [1], cuyo espacio de vectores de estado consta de $2^{|E|}$ elementos.

Ejemplo 2.2 Sea $G = (V, E)$ el grafo puente de la Figura 2.2, con nodos objetivo $K = \{s, t\}$ y espacio de vectores de estado Ω_G con $2^{|E|} = 2^5$ elementos. La confiabilidad s - t -terminal es la suma de la probabilidad de cada uno de los vectores de estado operativo posibles:

$$\begin{aligned} R_{\{s,t\}} = & r_{t,v}r_{t,u}r_{v,u}r_{v,s}r_{u,s} + q_{t,v}r_{t,u}r_{v,u}r_{v,s}r_{u,s} + r_{t,v}q_{t,u}r_{v,u}r_{v,s}r_{u,s} + \\ & r_{t,v}r_{t,u}q_{v,u}r_{v,s}r_{u,s} + r_{t,v}r_{t,u}r_{v,u}q_{v,s}r_{u,s} + r_{t,v}r_{t,u}r_{v,u}r_{v,s}q_{u,s} + \\ & q_{t,v}r_{t,u}q_{v,u}r_{v,s}r_{u,s} + q_{t,v}r_{t,u}r_{v,u}q_{v,s}r_{u,s} + q_{t,v}r_{t,u}r_{v,u}r_{v,s}q_{u,s} + \\ & r_{t,v}q_{t,u}q_{v,u}r_{v,s}r_{u,s} + r_{t,v}q_{t,u}r_{v,u}q_{v,s}r_{u,s} + r_{t,v}q_{t,u}r_{v,u}r_{v,s}q_{u,s} + \\ & r_{t,v}r_{t,u}q_{v,u}q_{v,s}r_{u,s} + r_{t,v}r_{t,u}q_{v,u}r_{v,s}q_{u,s} + q_{t,v}r_{t,u}q_{v,u}q_{v,s}r_{u,s} + \\ & r_{t,v}q_{t,u}q_{v,u}r_{v,s}q_{u,s} \end{aligned}$$

Considerando la igualdad $q_k = 1 - r_k$ la confiabilidad puede también escribirse:

$$\begin{aligned} R_{\{s,t\}} = & 2r_{u,s}r_{t,u}r_{v,u}r_{v,s}r_{u,s} - r_{u,s}r_{t,u}r_{v,u}r_{v,s} - r_{u,s}r_{t,u}r_{v,u}r_{u,s} - \\ & r_{u,s}r_{t,u}r_{v,s}r_{u,s} - r_{u,s}r_{v,u}r_{v,s}r_{u,s} - r_{t,u}r_{v,u}r_{v,s}r_{u,s} + r_{u,s}r_{u,t}r_{u,v} + \\ & r_{v,s}r_{v,t}r_{u,v} + r_{u,s}r_{u,t} + r_{v,s}r_{v,t} \end{aligned}$$

2.4.2.2. Principio de inclusión y exclusión y Método de Poincaré

La aplicación del principio de inclusión y exclusión, y su generalización, la fórmula de Poincaré, permite el cálculo exacto de la medida de confiabilidad en tiempo exponencial en el tamaño de la red. Aunque su complejidad hace prohibitivo el uso generalizado, la aplicación para redes pequeñas es válida.

El método requiere calcular la probabilidad de operatividad de cada uno de los *pathsets*, con lo cual la obtención del resultado es aún más costoso que la enumeración completa de vectores de estado. La fórmula contiene $2^P - 1$ términos y P , el número de *pathsets* es de orden exponencial en el número de enlaces de la red [4].

De acuerdo con la propiedad de inclusión-exclusión se cumple

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B).$$

La generalización de Poincaré extiende la fórmula J eventos, $J \geq 2$, forma que $\Pr(\bigcup_{j=1}^J A_j)$ es igual a:

$$\sum_j \Pr(A_j) - \sum_{j < i} \Pr(A_i A_j) + \sum_{j < i < l} \Pr(A_j A_i A_l) - \dots + (-1)^{J+1} \Pr(A_1 A_2 A_3 \dots A_J). \quad (2.8)$$

Si P_i^O es el i -ésimo pathset del grafo y la medida de confiabilidad es la probabilidad de que al menos un pathset esté operativo entonces de acuerdo con la generalización 2.8 la medida de confiabilidad de la red puede expresarse como:

$$R_K(G) = \sum_i \Pr(P_i^O) - \sum_{i < j} \Pr(P_i^O P_j^O) + \sum_{i < j < l} \Pr(P_i^O P_j^O P_l^O) - \dots + (-1)^{k+1} \Pr(P_1^O P_2^O P_3^O \dots P_k^O).$$

$$R_K(G) = \sum_{m=1}^P (-1)^{m-1} \sum_{1 \leq h_1 \leq h_2 \leq \dots \leq h_m \leq P} \Pr\left(\bigcap_{j=1}^m P_{h_j}^O\right),$$

De acuerdo con el modelo asumido, el funcionamiento de los enlaces es independiente. La expresión final para la medida es entonces:

$$R_K(G) = \sum_{m=1}^P (-1)^{m-1} \sum_{1 \leq h_1 \leq h_2 \leq \dots \leq h_m \leq P} \prod_{i \in \bigcup_{j=1}^m P_{h_j}^O} r_i. \quad (2.9)$$

Satyanarayana-Prabhakar [17] proponen una fórmula basada en conceptos de dominancia, que permite reducir el esfuerzo computacional respecto del método de Poincaré.

Definición 2.4.1 Una i -**formación** de un grafo G es un conjunto de i minpaths cuya unión es G . Si i es par (impar) la formación es par (impar).

Definición 2.4.2 La “signed domination” de un grafo G es el número de formaciones pares de G menos el número de formaciones impares de G , y se denota como $sdom(G)$.

La confiabilidad queda definida por: $R_K(G) = \sum_{H \subseteq G} sdom(H) \Pr(H)$, donde H varía en todos los vectores de estado de G y $\Pr(H)$ es la probabilidad de ocurrencia de H . La enumeración completa de configuraciones de la red se sustituye por el cálculo de $sdom$ de cada configuración de la red.

2.4.2.3. Métodos de Factorización

El método exacto de factorización, definido en el año 1958 por Moskowitz [18], propone una fórmula recursiva aplicada sobre subgrafos del grafo original, y cuya topología se obtiene de condicionar G al estado de un enlace e . El cálculo exacto de la medida de confiabilidad $R_K(G)$, para cualquier enlace e se basa en la siguiente descomposición pivotal:

$$R_K(G) = r_e R_{K'}(G * e) + (1 - r_e) R_K(G - e), \quad (2.10)$$

donde e se denomina arista pivot. La confiabilidad elemental de e es r_e , $G * e$ y $G - e$ son los grafos que resultan de G luego de la contracción (Procedimiento 1, Subsección 3.3.1) o eliminación (Procedimiento 2, Subsección 3.3.1) de e respectivamente. El conjunto K' coincide con K si al menos uno de los nodos extremos de la arista contraída no está en K . Si ambos

extremos pertenecen a K entonces K' sustituye los nodos por un único nodo que resulta de fusionarlos: $K' = (K - \{u, v\}) \cup z$.

La relación 2.10 se repite en cada una de las subredes obtenidas, y, en el peor de los casos, el número de repeticiones corresponde con la enumeración de todos los componentes de la red requiriendo tiempo de cómputo exponencial. El volumen de memoria de almacenamiento requerido es a lo sumo cuadrático en el tamaño de la red, necesitando almacenar como máximo $|E|$ llamadas recursivas [19]. El mecanismo de selección de la arista pivot influye en el desempeño del algoritmo.

En el caso de confiabilidad s - t -terminal se continúa hasta que al eliminar una arista no existan caminos entre s y t o cuando por la contracción de un enlace los nodos terminales quedan conectados de forma que la red está siempre operativa.

El Algoritmo 1 resume el método de factorización, basándose en dos procedimientos auxiliares, $Reduce(G)$ y $Select_{pivot}(G)$.

ALGORITMO 1 FACTORIZACIÓN

ENTRADA:

grafo G ,
 nodos objetivo K ,
 confiabilidad de los enlaces \vec{r} ,
 número de iteraciones N .

SALIDA:

medida $R_K(G)$.

- 1: Aplicar $Reduce(G)$.
 - 2: **Si** No es conexo $G = (V, E)$ **entonces**
 - 3: $R_K(G) = 0$
 - 4: **Fin Si**
 - 5: **Si** $|K| = 1$ **entonces**
 - 6: $R_K(G) = 1$
 - 7: **Fin Si**
 - 8: **Si** $(G_K == e)$ y $(|K| = 2)$ **entonces**
 - 9: r_e
 - 10: **Si no**
 - 11: $e = Select_{pivot}(G)$
 - 12: $R_K(G) = r_e FACTO(G_k * e) + q_e FACTO(G_K - e)$
 - 13: **Fin Si**
-

$Reduce(G)$ aplica procedimientos de reducción sobre el grafo G con el objetivo de simplificar la topología y reducir el número de operaciones necesarias (ver Subsección 3.3.1).

$Select_{pivot}(G)$ designa el procedimiento de selección de la arista pivot. Algunas consideraciones pueden consultarse en [19].

Los parámetros de entrada del Algoritmo 1 son el grafo $G = (V, E)$ con nodos objetivo K y

confiabilidad elemental de los enlaces r . Retorna el valor exacto de la medida de confiabilidad $R_K(G)$.

En la Línea 1 se aplican procedimientos de reducción a G , obteniendo un grafo equivalente en términos de la medida de confiabilidad.

En el bloque comprendido entre las Líneas 2 y 10 se verifican casos base de la recursión que corresponden a grafos de topología trivial. En caso de cumplirse alguna de las 3 condiciones indicadas en las Líneas 2, 5 y 8 la búsqueda recursiva finaliza.

En la Línea 2 se evalúa si el grafo está en estado de falla. Si ésta condición se cumple la medida de confiabilidad $R_K(G) = 0$.

En la Línea 5 se verifica si el conjunto objetivo K tiene un único nodo. En caso de cumplirse la condición el grafo siempre está operativo y por lo tanto $R_K(G) = 1$.

En la Línea 8 se evalúa si la red consta de un único enlace. La probabilidad de que la red esté operativa es r_i y el Algoritmo 1 retorna $R_K(G) = r_i$.

Si el grafo no corresponde a alguno de los casos previos se aplica la recursión 2.10.

En la Línea 11 se selecciona la arista pivot a partir de la cual se definen los subgrafos de G , $G * e$ y $G - e$.

Por último en la Línea 12 se aplica la Ecuación 2.10, y el proceso avanza hasta que alguno de los subgrafos definidos coincide con alguno de los casos base.

Para finalizar se presenta un ejemplo tomado de [19].

Ejemplo 2.3 *Considere el grafo presentado en la Figura 2.8. El procedimiento de factorización parte del grafo G_0 y elige una arista pivot, $e_{4,5} = \{v_4, v_5\}$ a partir de la cual se desarrolla el proceso recursivo. La contracción de $e_{4,5}$ da lugar al grafo G_1 : $G_1 = G_0 * e_{4,5}$, y la eliminación de la misma da lugar a G_2 : $G_2 = G_0 - e_{4,5}$.*

En el grafo G_2 puede aplicarse reducción en serie sobre el camino v_1, v_4, v_3 , pues v_4 es de grado 2 y no es un nodo objetivo. De dicha reducción resulta un grafo que admite reducción paralelo a las dos aristas que conectan v_3 y v_1 ; en forma análoga se procede con v_1, v_5, v_2 .

Se obtiene un camino de largo 2 entre v_3 y v_2 al que nuevamente se aplica reducción en serie dando lugar a $G_3 = e_{2,3}$ (Ver Figura 2.9).

El grafo resultante de la contracción inicial, G_1 , admite reducción paralelo de las aristas que unen v_1 y v_4 . Si se toma la arista resultante de dicha transformación como pivot, se obtienen G_4 por contracción y G_5 por eliminación.

De los grafos resultantes, G_4 admite 2 reducciones arista paralela, generando un camino de largo 2, $\{v_3, v_1, v_2\}$ que por reducción en serie genera $G_6 = e_{2,3}$.

El grafo G_5 admite doble reducción en serie, para los caminos v_2, v_4, v_3 y v_2, v_1, v_3 . El grafo resultante consta de dos nodos con dos aristas en paralelo, al que luego de una reducción, da lugar al grafo $G_7 = e_{2,3}$.

Los grafos G_3 , G_6 y G_7 coinciden con unos de los casos base del Algoritmo 1, $G = e$, y la búsqueda recursiva finaliza.

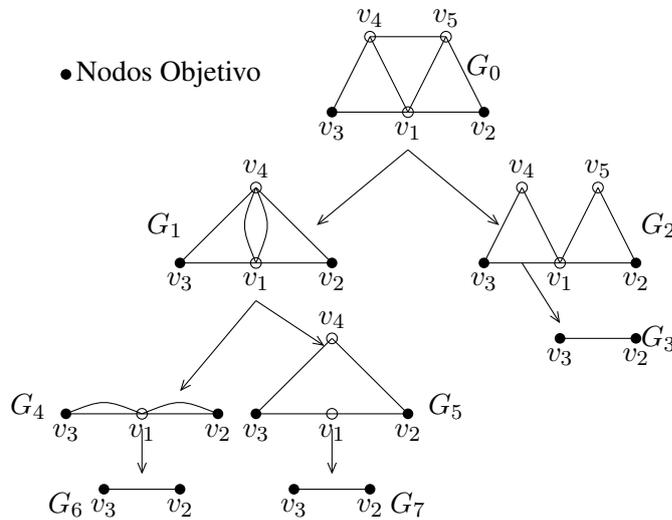
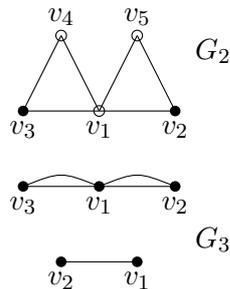


Figura 2.8: Ejemplo Algoritmo Factorización

Figura 2.9: Reducción del subgrafo G_2

2.4.3. Métodos aproximados

El esfuerzo computacional requerido para calcular la confiabilidad depende del tamaño de la red que se desea evaluar. Ahora bien, en general ¿cuántos nodos tiene una red real? En trabajos recientes Provan [20] argumenta que para considerar una red grande, desde el punto de vista de la evaluación de la confiabilidad, el número de nodos debe situarse en torno a 50. En trabajos anteriores se considera que una red de más de 20 nodos hace el cálculo impracticable [21]. Por otra parte, en el área de comunicaciones el tamaño es usualmente más alto, aunque de acuerdo con la información publicada la dimensión de una red real frecuentemente no supera los 200 nodos [22].

La importancia de la confiabilidad y el tamaño de las redes reales junto con la complejidad computacional asociada al análisis de la $R_K(G)$ medida ha motivado el continuo esfuerzo por

encontrar métodos de resolución más rápidos. Los métodos exactos no tienen buen desempeño en redes de más de 50 nodos y el tiempo de cálculo crece en forma exponencial con el tamaño de la red [4].

En el caso general el problema pertenece a la clase $\#P$ -completo, clase que pertenece a una familia de problemas NP -difícil no conocida que sea NP [1, 4]. Un problema $\#P$ -completo es equivalente a contar el número de soluciones de uno NP -completo, y por lo tanto un problema $\#P$ -completo es al menos tan difícil como uno NP -completo.

Incluso para el caso de $K = \{s, t\}$ el problema de evaluar la confiabilidad de manera exacta es NP -difícil [6].

Además de la complejidad asociada a la dimensión de las redes actuales, otra característica a tener en cuenta es la baja probabilidad de falla de sus componentes. Gracias al desarrollo de las redes y su alta densidad los sistemas actuales son altamente confiables y la falla de sus componentes es un evento de muy baja probabilidad, *eventos raros* [23]. En este contexto, métodos aproximados de amplia aplicación como el método Monte Carlo Estándar (*CMC* por sus siglas en inglés, Crude Monte Carlo) no obtienen buen desempeño. La varianza del estimador crece a medida que el evento de falla se vuelve más raro requiriendo un número excesivo de iteraciones para generar estimaciones de buena calidad [23].

Para corregir la dependencia entre la precisión de la estimación y el valor de la confiabilidad elemental de los enlaces surgen métodos de reducción de varianza que ofrecen estimaciones más precisas, aunque son costosos en términos del tiempo y memoria.

En esta sección se presenta el método *CMC* aplicado para estimar la medida de confiabilidad $R_K(G)$ y un método alternativo que permite obtener mejoras en el costo computacional manteniendo la precisión. En las subsecciones siguientes se detallan métodos de reducción de varianza, y que como indica su nombre, permiten obtener resultados con menos varianza que la obtenida por *CMC*.

2.4.3.1. Método Monte Carlo

El método Monte Carlo Estándar (*CMC*) [24] es relativamente sencillo, se adapta a diversos problemas y permite la estimación de parámetros en redes de gran porte con buen desempeño. En el problema de confiabilidad, la técnica genera N réplicas aleatorias independientes de configuraciones de la red, $X^{(i)}$ de G , con $i = 1, 2, \dots, N$. El número de iteraciones se define en función de la precisión que se desea alcanzar y del tiempo de cálculo máximo admitido. Cada componente del vector de estado $X^{(i)}$ se genera sorteando una variable aleatoria continua U de distribución de probabilidad Uniforme(0,1). Si U es menor o igual a la probabilidad de que el j -ésimo enlace esté operativo, r_j , entonces el enlace se fija en estado operativo. En caso contrario se establece en estado de falla. Formalmente, este concepto se resume en la función indicatriz $X_j = 1_{\{U \leq r_j\}}$.

El estimador de la medida de confiabilidad se define como la proporción de estados operativos sorteados en relación al total de sorteos realizados, y puede interpretarse como una tasa de operatividad de la red. En términos de $\Phi(X^{(i)})$ la confiabilidad K -terminal estimada $\hat{R}_K(G)$,

su varianza $V(R_K(G))$ y su estimación \hat{V} se definen respectivamente como:

$$\hat{R}_K(G) = \frac{\sum_{i=1}^N \Phi(X^{(i)})}{N}, \quad (2.11)$$

$$V(\hat{R}_K(G)) = \frac{R_K(G)(1 - R_K(G))}{N}, \quad (2.12)$$

$$\hat{V} = \frac{\hat{R}_K(G)(1 - \hat{R}_K(G))}{N - 1}. \quad (2.13)$$

ALGORITMO 2 Monte Carlo Crudo (CMC)

ENTRADA:

grafo G ,
 nodos objetivo K ,
 confiabilidad de los enlaces \vec{r} ,
 número de iteraciones N .

SALIDA:

estimación de la medida: $\hat{R}_K(G)$,
 estimación de la varianza: \hat{V} .

- 1: $Y = 0$
 - 2: **Para** $i = 1$ **hasta** N
 - 3: $x_e^{(i)} = 0, \forall e = 1, 2, \dots, m$
 - 4: **Para** $j = 1$ **hasta** m
 - 5: Sorteo $U \sim Uniforme(0, 1)$.
 - 6: **Si** $U \leq r_j$ **entonces**
 - 7: $x_j^{(i)} = 1$
 - 8: **Fin Si**
 - 9: $Y = Y + \Phi(x^{(i)})$
 - 10: **Fin Para**
 - 11: **Fin Para**
 - 12: $\hat{R}_K(G) = \frac{Y}{N}$
 - 13: $\hat{V} = \frac{\hat{R}_K(G)(1 - \hat{R}_K(G))}{N - 1}$
-

El tiempo de cómputo del método *CMC* es lineal en el número de aristas del grafo. El muestreo del estado de cada una de las aristas y la verificación del estado global del grafo tiene una complejidad de orden $O(|E|)$, como se efectúan N iteraciones la complejidad global es entonces $O(N|E|)$ [23].

En el Algoritmo 2 los parámetros de entrada son el grafo $G = (V, E)$, el conjunto de nodos objetivo K , la confiabilidad elemental de sus enlaces $\vec{r} = (r_1, r_2, \dots, r_{|E|})$ y el número de iteraciones N . Los parámetros de salida son la estimación de la medida de confiabilidad $R_K(G)$ y su varianza.

En la Línea 1 del algoritmo se inicializa una variable de conteo Y , en la que se almacena el número de estados operativos de G sorteados.

En la Línea 2 se indica el número de iteraciones que coincide con el tamaño de muestra de estados de la red G .

En el bloque de código comprendido entre las Líneas 3 y 10 se sortea un vector de estado de G .

En la Línea 4 se inicializan todos los enlaces en estado de falla, y en la Línea 5 se sortea una variable aleatoria U Uniforme(0,1) a partir de la que se define el estado de cada enlace: si U es menor o igual que la confiabilidad elemental de la arista j -ésima entonces se fija en estado operativo, $x_j^{(i)} = 1$.

En la Línea 9, se incrementa el contador Y en $\Phi(x^{(i)})$. Si el estado sorteado es operativo Y aumenta en una unidad, de lo contrario permanece igual ($\Phi(x^{(i)}) = 0$).

En la Línea 12 se obtiene la estimación de la confiabilidad definida como la proporción de estados operativos sorteados respecto al número de iteraciones, y en la Línea 14 se calcula la varianza, según Ecuación 2.13.

El desempeño del método *CMC* no es robusto en redes altamente confiables, donde la ocurrencia de fallas es un evento raro [23]. Independientemente del tamaño de la red el error relativo no está acotado superiormente cuando la anticonfiabilidad es muy pequeña:

$$ER \rightarrow \infty \quad \text{cuando} \quad \tilde{Q} \rightarrow 0.$$

Definición 2.4.3 Sea σ_n^2 la varianza de un estimador $\hat{\delta}$ y z_α el valor asociado a la distribución normal estándar Φ tal que $z_\alpha = \Phi^{-1}(1 - \alpha)$. El **error relativo (ER)** asociado al estimador $\hat{\delta}$ es la mitad del intervalo de confianza:

$$ER = z_\alpha \frac{\sigma_n}{\hat{\delta}}. \quad (2.14)$$

En el método *CMC* se cumple

$$ER = z_\alpha \frac{\sigma_n}{\hat{Q}} \approx \frac{1}{(N\hat{Q})^{\frac{1}{2}}},$$

donde \hat{Q} es la estimación de la medida de anticonfiabilidad y $\sigma_n = \sqrt{\frac{\hat{Q}(1-\hat{Q})}{N}}$ [23].

2.4.3.2. Formulación alternativa de Monte Carlo Crudo

La reformulación del método *CMC*, propuesta por El Khadiri y Rubino [25], mejora el desempeño computacional de *CMC* alcanzando la misma precisión:

$$V_{AMC} = V_{CMC},$$

$$T_{AMC} \leq T_{CMC},$$

donde V_{AMC} , T_{AMC} y V_{CMC} , T_{CMC} son la varianza y el tiempo de cálculo de la formulación alternativa y del método *CMC* respectivamente.

Dada una red $G = (V, E)$ el método requiere identificar un conjunto de minpaths arista disjuntos de G , $P = \{P_1, \dots, P_H\}$. Bajo los supuestos del modelo probabilístico de confiabilidad la falla de los enlaces ocurre en forma independiente con lo cual la probabilidad de que un minpath P_h opere es

$$\Pr(\text{"}P_h \text{ opera"}) = p_h = \prod_{e_i \in P_h} r_i, \quad (2.15)$$

mientras que la probabilidad de que al menos una arista del minpath falle es el complemento de la Ecuación 2.15, $q_h = 1 - p_h$.

Como los minpaths de P son arista disjuntos, la probabilidad de que al menos una arista de cada minpath falle es:

$$q = \prod_{h=1}^H q_h, \quad (2.16)$$

condición necesaria pero no suficiente para que la red no esté operativa.

Asociado a la probabilidad de falla q se define la variable F como:

$$F = \min\{j \geq 1 : \Phi(x^{(j)}) = 0\}$$

siendo $x^{(1)}, x^{(2)}, \dots, x^{(N)}$ una muestra de vectores de estado de G .

En términos de una muestra probabilística F puede interpretarse como el número de estados de la red que se deben sortear hasta identificar un estado de falla, lo que permite asociar a F una distribución de probabilidad Geométrica de parámetro q .

Definición 2.4.4 Sea X una variable aleatoria Geométrica de parámetro p , donde p es la probabilidad de éxito de sorteos independientes Bernoulli. La probabilidad de realizar x intentos hasta obtener un éxito es:

$$\Pr(X = x) = (1 - p)^{x-1} p^x, x \in N^*.$$

La variable X representa el número de sorteos necesarios para obtener un éxito. Su valor esperado y su varianza son respectivamente $E(X) = \frac{1}{p}$ y $Var(X) = \frac{1-p}{p^2}$.

Se cumple entonces:

$$\Pr(F = n) = (1 - q)^{n-1} q,$$

donde q es la probabilidad de ocurrencia de una falla y n es el número de intentos necesarios hasta generar un sistema en estado de falla.

El valor esperado y la varianza de F son $\frac{1}{q}$ y $\frac{1-q}{q^2}$. Nótese que cuanto más cercano a 0 es q menor es la probabilidad de que el sistema deje de estar operativo y mayor es el número de muestras necesarias para generar un estado de falla.

La reformulación de CMC permite disminuir el costo computacional en el factor q : $Nq|E|$ [25]. (En el método clásico el costo es $N|E|$).

ALGORITMO 3 Reformulación de CMC**ENTRADA:**

grafo G ,
 nodos objetivo K ,
 confiabilidad de los enlaces \vec{r} ,
 número de iteraciones N .

SALIDA:

estimación de la anticonfiabilidad de la red: $\hat{Q}_K(G)$,
 estimación de la varianza: $Var(\hat{Q}_K(G))$.

```

1:  $Y = 0$ 
2:  $F = Geo(q)$ 
3: Mientras  $F < N$ 
4:    $g = MuestraCondicional(G)$ 
5:   Si  $g$  es un estado de falla entonces
6:      $Y = Y + 1$ 
7:   Fin Si
8:    $F = F + Geo(q)$ 
9: Fin Mientras
10:  $\hat{Q}_K(G) = \frac{Y}{N}$ 
11:  $Var(\hat{Q}_K(G)) = \hat{V} = \frac{\hat{Q}_K(G)(1-\hat{Q}_K(G))}{N-1}$ 

```

En el Algoritmo 3 los parámetros de entrada son el grafo $G = (V, E)$, el conjunto de nodos objetivo K , la confiabilidad elemental de sus enlaces \vec{r} y el número de iteraciones N . Los parámetros de salida son la estimación de la medida de anticonfiabilidad $Q_K(G)$ y su varianza. En el método se aplica muestreo condicional sobre el espacio de vectores de estado de G , de forma que se consideran sólo aquellas configuraciones en que todos los minpaths tienen al menos un enlace no operativo.

En la Línea 1 se define una variable de conteo Y , que se inicializa en 0 y aumenta en 1 cada vez que se identifica un estado de falla.

En la Línea 2 se sortea una variable aleatoria F con distribución Geométrica de parámetro q , donde q es la probabilidad definida en la Ecuación 2.16. F representa el número de estados de G que se deben sortear para generar un estado de falla. Si F es menor que N se simula el estado de la red y se evalúa si falla. Si en cambio el valor de F supera N significa que la probabilidad de éxito es muy pequeña y el valor de N fijado no es suficiente para identificar una falla del sistema.

En las Líneas 3 a 9 se realizan iteraciones mientras la suma del tamaño de las muestras sorteadas no totaliza N .

En la Línea 4 se obtiene una muestra condicional del grafo G , condicionada a que al menos una arista de cada minpath falle, con costo computacional lineal en el número de aristas. El muestreo condicional no garantiza que la red falle a menos que los minpaths sean independientes.

Si el estado sorteado g , es un estado de falla de la red el contador Y aumenta en una unidad como indica la Línea 6.

En la Línea 8 se sortea otra variable Geométrica, que indica cuántos estados de G se deben sortear para identificar un nuevo estado de falla.

En [26] se propone un método que, además de reducir el tiempo de cálculo como propone la formulación alternativa a CMC , reduce la varianza reemplazando la variable aleatoria geométrica por su valor esperado. En la siguiente sección se consideran métodos generales de reducción de varianza.

2.4.3.3. Métodos de reducción de varianza

En los métodos de estimación es de interés obtener resultados precisos y a bajo costo computacional. Si existen dos métodos, M_1 y M_2 , que estiman sin sesgo una medida de interés R con varianzas \hat{V}_1 y \hat{V}_2 entonces el método M_2 es preferible a M_1 si $\hat{V}_2 < \hat{V}_1$. La reducción de la varianza genera intervalos de confianza de menor amplitud, y aunque esta reducción puede lograrse incrementando el tamaño de muestra, no necesariamente implica incremento en el tiempo de cómputo [23].

Esta sección enumera métodos generales de reducción de varianza que aseguran disminución del error de estimación respecto del método CMC para un mismo tamaño de muestra y con un esquema de funcionamiento similar.

En particular se presentan los siguientes métodos: Muestreo Basado en Cotas, Método Antitético, Muestreo Dagger, Método de Construcción-Reducción Secuenciales, Método de Riesgo Total, Muestreo de Importancia, Método de Cubrimiento y Método RVR .

Muestreo Basado en Cotas

La primera aplicación del método basado en cotas en el contexto de confiabilidad en redes fue desarrollada por Van Slyke y Frank [15] en el año 1971. Posteriormente realizaron publicaciones Kumamoto (1977) [27] y Fishman (1986) [28].

A partir de un grafo $G = (V, E)$, configuración X y espacio de vectores de estado Ω_G , el método propone realizar muestreo en un sub-espacio de Ω_G .

La aplicación de muestreo basado en cotas considera dos funciones Φ^L y Φ^U que acotan la función de estructura Φ en ambos extremos cumpliéndose $\Phi^L(x) \leq \Phi(x) \leq \Phi^U(x)$, $\forall x \in \Omega_G$.

Se definen además la probabilidad de que cada cota valga 1, condicionada al estado de k aristas:

- $R_k^L(\tilde{X}^{(k)}) = \Pr(\Phi^L(X) = 1 | X_1 = \tilde{x}_1, \dots, X_k = \tilde{x}_k)$.
- $R_k^U(\tilde{X}^{(k)}) = \Pr(\Phi^U(X) = 1 | X_1 = \tilde{x}_1, \dots, X_k = \tilde{x}_k)$.

donde $k = 0, \dots, |E|$, $\tilde{X}^{(k)} = (\tilde{X}_1, \dots, \tilde{X}_k)$.

La probabilidad de que las cotas valgan 1 es computable en tiempo polinomial [29] y en el

caso particular $k = 0$ los valores R_0^L y R_0^U cumplen:

$$\begin{aligned} R_0^L &= \Pr(\Phi^L(X) = 1), \\ R_0^U &= \Pr(\Phi^U(X) = 1). \end{aligned}$$

Se define un espacio W conformado por los vectores de estado de G tal que la cota inferior vale 0 y la superior vale 1:

$$W = \{X : \Phi^L(X) = 0, \Phi^U(X) = 1\}.$$

La muestra se obtiene de acuerdo con una función de probabilidad proporcional a la correspondiente en el espacio original Ω_G . En particular cada componente \tilde{X}_i del vector de estados $\tilde{X} \in \mathbb{R}^{|E|}$ se sortea con probabilidad:

$$\tilde{p}_i = \Pr(X_i = 1 | X_1 = \tilde{x}_1, \dots, X_{i-1} = \tilde{x}_{i-1}, \Phi^U(\tilde{X}) = 1, \Phi^L(\tilde{X}) = 0)$$

$$\tilde{p}_i = p_i \cdot \left[\frac{R_i^U(\tilde{X}^{i-1}) - R_i^L(\tilde{X}^{i-1})}{R_{i-1}^U(\tilde{X}^{i-1}) - R_{i-1}^L(\tilde{X}^{i-1})} \right].$$

Dada la muestra se obtiene un estimador de la confiabilidad del sistema según:

$$\hat{R}_K(G) = R_0^L + \Phi(\tilde{X})(R_0^U - R_0^L). \quad (2.17)$$

ALGORITMO 4 Muestreo basado en cotas

ENTRADA:

grafo G ,
 nodos objetivo K ,
 cota inferior R_0^L ,
 cota superior R_0^U .

SALIDA:

estimación de la medida de confiabilidad: $\hat{R}_K(G)$

- 1: **Para** $i = 1$ **hasta** m
 - 2: $\tilde{X}_i = \text{Ber}(\tilde{p}_i)$
 - 3: **Fin Para**
 - 4: $\hat{R}_K(G) = R_0^L + \Phi(\tilde{X})(R_0^U - R_0^L)$
-

El Algoritmo 4, a partir del grafo $G = (V, E)$, el conjunto de nodos objetivo K y los valores R_0^L, R_0^U estima la medida de confiabilidad $R_K(G)$.

En las Líneas 1 a 3 se genera el vector \tilde{X} sorteando para cada enlace una variable aleatoria Bernoulli, tal que la probabilidad de éxito del i -ésimo enlace es \tilde{p}_i .

Dada la muestra, en la Línea 4 se estima el valor de la medida $R_K(G)$ de acuerdo con la Ecuación 2.17.

La reducción de la varianza es directamente proporcional a la parte de la probabilidad que

queda incluida en el espacio W y la magnitud de la disminución depende del grado de precisión de las cotas obtenidas [29]. Mientras para el método *CMC* la varianza de $\hat{R}_K(G)$ es $R_K(G)(1 - R_K(G))$, la varianza obtenida con muestreo basado en cotas es [29]:

$$R_K(G)(1 - R_K(G)) - (1 - R_0^U)R_K(G) - R_0^L(R_0^U - R_K(G)).$$

Método antitético

En el método antitético la reducción de varianza se obtiene a partir del sorteo “guiado” de variables aleatorias uniformes en las que se introduce correlación negativa.

La idea central surge de la definición de varianza de la suma de dos variables aleatorias X, Y :

$$Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y).$$

Nótese que si X, Y son independientes $Cov(X, Y) = 0$, mientras que si la correlación $\rho_{X,Y}$ es negativa también lo es la covarianza. La correlación entre X e Y se define:

$$\rho_{X,Y} = \frac{Cov(X, Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}.$$

Como $\sqrt{Var(\cdot)}$ es siempre positiva, si $\rho_{X,Y} < 0$ se cumple $Cov(X, Y) < 0$.

Por otro lado el estimador de la medida de confiabilidad se obtiene como el promedio de N realizaciones, $\hat{R}_K(G) = \frac{\sum_{i=1}^N \Phi(x^{(i)})}{N}$, y su varianza es:

$$Var(\hat{R}_K(G)) = \frac{\sum_{i=1}^N Var(\Phi(X^{(i)}))}{N^2} + \frac{2}{N^2} \sum_{i < j} Cov(\Phi(X^{(i)}), \Phi(X^{(j)})). \quad (2.18)$$

Si las réplicas para estimar la medida $R_K(G)$ tienen correlación negativa, la varianza del método antitético es menor que la varianza de *CMC* donde las realizaciones son independientes.

Una forma de selección que asegura correlación negativa entre las variables es sortear variables aleatorias independientes de distribución uniforme y aplicar el siguiente criterio:

$$\begin{cases} X^{(1)} = 1_{\{U < r\}} \\ X^{(2)} = 1_{\{1-U < r\}} \end{cases} \quad (2.19)$$

Proposición 2.4.5 Sean $X^{(1)}$ y $X^{(2)}$ las variables definidas en la Ecuación 2.19, entonces:

$$Cov(X^{(1)}, X^{(2)}) = -q^2,$$

donde

$$q = \begin{cases} r, & \text{Si } r < 0.5, \\ 1 - r, & \text{Si } r \geq 0.5. \end{cases} \quad (2.20)$$

Demostración.

$$E(X^{(1)}) = \Pr(U < r) = r, \quad (2.21)$$

$$E(X^{(2)}) = \Pr(U > 1 - r) = 1 - \Pr(U < 1 - r) = 1 - (1 - r) = r. \quad (2.22)$$

En el caso general la covarianza se define como:

$$Cov(X^{(1)}, X^{(2)}) = E(X^{(1)}X^{(2)}) - E(X^{(1)})E(X^{(2)}).$$

Si $r < 0.5$ se cumple:

$$Cov(X^{(1)}, X^{(2)}) = E(X^{(1)}, X^{(2)}) - r^2 = 0 - r^2 = -r^2.$$

Por otro lado si $r \geq 0.5$:

$$Cov(X^{(1)}, X^{(2)}) = E(X^{(1)}, X^{(2)}) - r^2 = 2r - 1 - r^2 = -(r^2 - 2r + 1) = -(1 - r)^2.$$

De acuerdo con la definición de q en el caso general $Cov(X^{(1)}, X^{(2)}) = -q^2$

QED

En los modelos coherentes, si se cumple $Cov(X_e^{(1)}, X_e^{(2)}) < 0$ para algún enlace de G , según el Teorema 2.4.6 se cumple también para la función Φ , $Cov(\Phi(X^{(1)}), \Phi(X^{(2)})) < 0$ [30], siendo $X^{(1)}$ y $X^{(2)}$ dos vectores de estado de G . De acuerdo con la relación 2.18 se logra la reducción de varianza respecto de CMC .

Teorema 2.4.6 *Sea Φ una función de estructura coherente. Si para cada arista $e \in E$ las variables $X_e^{(i)}$ y $X_e^{(j)}$ cumplen $Cov(X_e^{(i)}, X_e^{(j)}) \leq 0$ y al menos uno de los enlaces cumple $Cov(X_e^{(i)}, X_e^{(j)}) < 0$ entonces $Cov(\Phi(X^{(i)}), \Phi(X^{(j)})) < 0$.*

El Cuadro 2.1 resume la probabilidad de ocurrencia del par $X^{(1)}$ y $X^{(2)}$ en función de los valores que toma U [30]. Dicha probabilidad coincide con la amplitud del intervalo en el que se evalúa la probabilidad. Por ejemplo $\Pr(U \in [1 - r, 1)) = r$ y la amplitud de $[1 - r, 1)$ es $1 - (1 - r) = r$.

Cuadro 2.1: Probabilidad conjunta de las variables antitéticas

$r < 0.5$			$r \geq 0.5$		
$X^{(1)}$	$X^{(2)}$	Prob	$X^{(1)}$	$X^{(2)}$	Prob
1	0	$\Pr(U < r) = r$	1	0	$\Pr(U \in [0, 1 - r)) = 1 - r$
0	1	$\Pr(U \in [1 - r, 1)) = r$	0	1	$\Pr(U \in [r, 1)) = 1 - r$
0	0	$\Pr(U \in [r, 1 - r)) = 1 - 2r$	0	0	0
1	1	0	1	1	$\Pr(U \in [1 - r, r)) = 2r - 1$

Dentro de los métodos basados en variables antitéticas en [31] se propone una variante para redes donde la confiabilidad elemental es idéntica para todos los enlaces. En particular se calcula una cota inferior R_O^L , evaluando la confiabilidad en subconjuntos de vectores de estado

tal que el número de enlaces no operativos es menor que el cardinal del mincut de G . En este caso el costo computacional es similar a CMC y la reducción en la varianza es del orden de $\frac{1}{1-R_O^L}$ [31]. **Muestreo Dagger**

Este método fue propuesto en [32] por Kumamoto et al y se clasifica dentro de los métodos de tipo antitético. Una forma sencilla de definir dos variables antitéticas es tomar el vector aleatorio $(X^{(1)}, X^{(2)})$, con $X^{(1)}$ y $X^{(2)}$ Bernoulli(r):

$$X^{(i)} = (1 - X)1_{U \in I^i} + X1_{U \notin I^i},$$

donde $X = 1_{\{r \geq 0.5\}}$. El método Dagger generaliza esta idea para M variables antitéticas. El parámetro M es el tamaño de muestra y se elige de forma que las M réplicas puedan agruparse en M_e submuestras de tamaño $N_e \forall e \in E$. El parámetro M se obtiene como el mínimo común múltiplo de N_e donde:

$$N_e = \left\lceil \frac{1}{q_e} \right\rceil,$$

donde q_e es la anticonfiabilidad elemental del enlace e . A cada enlace se asocian M_e bloques o intervalos de amplitud I^i , $i = 1, 2, \dots, M_e$ tal que $\bigcup_{i=1}^{M_e} I^i = [0, 1]$. Cuando $\bigcup_{i=1}^{M_e} I^i$ no cubre todo el intervalo $[0, 1]$ se define un intervalo remanente $J = [0, 1] - \bigcup_{i=1}^{M_e} I^i$.

Para cada arista e , en cada submuestra de tamaño N_e , se sortea una variable aleatoria Uniforme(0,1) y se define un parámetro $f = \lceil \frac{U}{q_e} \rceil$ tal que la arista e sólo falla en el elemento f de la submuestra de tamaño N_e . La forma de selección asegura que la frecuencia de falla en la arista es proporcional a su tasa media de falla q_e .

Ejemplo 2.4 Sea $G = (V, E)$, tal que $|E| = 3$, $r_1 = 0.5$, $r_2 = 0.75$, $r_3 = 0.84$ y $q_i = 1 - r_i$, $i = 1, 2, 3$. Se calcula N_e para cada enlace, el tamaño de cada sub-bloque y se define el número de intervalos $M_e = \frac{M}{N_e}$, con $M = \text{mcm}(N_1, N_2, N_3) = 12$. En el único caso que se obtiene un intervalo remanente J es en el tercer enlace. Los valores de cada parámetro se detallan en la Tabla 2.4.

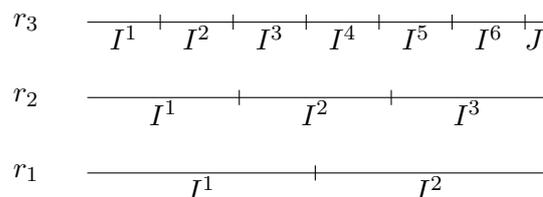


Figura 2.10: Intervalos de los enlaces.

Cuadro 2.2: Valores de los parámetros

e	q_e	N_e	$M_e = \frac{M}{N_e}$
1	0.5	2	6
2	0.25	4	3
3	0.16	6	2

ALGORITMO 5 Muestreo Dagger**ENTRADA:**

grafo G ,
 nodos objetivo K ,
 confiabilidad de los enlaces \vec{r} .

SALIDA:

estimación de la medida de confiabilidad: $\hat{R}_K(G)$

- 1: Calcular vector de enteros N tal que $N_e = \lfloor \frac{1}{q_e} \rfloor$.
- 2: Elegir el tamaño de muestra $M = mcm\{N_e : e \in E\}$
- 3: $Y = 0$
- 4: **Para** $e = 1$ **hasta** $|E|$
- 5: $|M_e| = \frac{M}{N_e}$.
- 6: **Para** $i = 1$ **hasta** M_e
- 7: Sortear U Uniforme.
- 8: $f = \lfloor \frac{U}{q_e} \rfloor$
- 9: **Si** $f \leq N_e$ **entonces**
- 10: $x_e^{(i)} = 0$
- 11: **Fin Si**
- 12: **Fin Para**
- 13: **Fin Para**
- 14: **Para** 1 **hasta** M
- 15: $Y = Y + \Phi(x^{(i)})$
- 16: **Fin Para**
- 17: $R_K(G) = \frac{Y}{M}$

El Algoritmo 5 resume el Método Dagger. Los parámetros de entrada son el grafo G , la confiabilidad de los enlaces, \vec{r} y el conjunto de nodos objetivo K . Retorna una estimación de la medida $R_K(G)$.

A partir del grafo G obtiene una estimación de la medida $R_K(G)$.

En la Línea 1 se calcula el parámetro N_e asociado a cada enlace $e \in E$.

En la Línea 2 se obtiene el tamaño de muestra M como el mínimo común múltiplo de los valores N_e .

En la Línea 3 se inicializa una variable de conteo del número de estados operativos de la red.

En las Líneas 4 y 5 se indican las iteraciones para cada enlace e y para cada subbloque de tamaño N_e .

El bloque entre las Líneas 6 y 11 se repite $|E| \times (\sum_{e \in E} M_e)$. La Línea 6 se obtienen los valores

M_e para todos los enlaces de G .

En la Línea 7 se fija el enlace e en estado operativo en todos los vectores en el bloque de tamaño N_e . Cada enlace cambia a estado de falla según una variable aleatoria Uniforme(0,1), indicado en la Línea 8, y el valor f .

En la Línea 9 se calcula el entero f , tal que si $f \leq N_e$ entonces e falla en el elemento f de la i -ésima submuestra de las M_e , como se indica en la Línea 10 y 11.

El procedimiento se repite para todos los enlaces de G . Por último en las Líneas 15-17 se evalúa cada uno de los M estados generados y se cuentan el número de estados operativos en Y . En la Línea 18 se obtiene el estimador de $R_K(G)$ como el porcentaje de estados operativos de G respecto del tamaño de muestra M .

Este método, si bien tiene el mismo tiempo de cálculo en cada muestra, genera un número menor de variables aleatorias lo que reduce el tiempo de ejecución respecto a *CMC* [29]. En cada bloque, la complejidad espacial es $O(E \times L)$

En [30] los autores señalan que una de las desventajas del método deriva de sus requerimientos de memoria, restringiendo su aplicación a algunos grafos. El mecanismo de selección asegura que la frecuencia de falla de cada enlace es directamente proporcional a su tasa media de falla y aleatoriza la forma en la que se combinan las fallas individuales para definir un estado de la red.

Método de construcción/reducción secuenciales

El método de construcción/reducción secuenciales considera un modelo en el que se asume que el estado de los enlaces de la red puede cambiar en el tiempo. En los métodos mencionados hasta el momento la red se asume estática.

Un ejemplo de modelos dinámicos son los métodos propuestos por Easton y Wong en 1980 [33] denominados métodos de construcción y destrucción secuencial.

En el caso de los métodos de construcción, se considera un punto de partida en el que todos los enlaces del sistema fallan, y en forma secuencial considerando un criterio de orden de los enlaces, se reparan hasta que el sistema se vuelve operativo.

Por el contrario, los métodos de destrucción inician el proceso con todos los enlaces operativos y, uno a uno, fija en estado de falla los enlaces hasta que el sistema falla.

El procedimiento logra reducir la varianza respecto de *CMC*, pero el costo computacional es mayor, del orden $O(|E| \max\{|V|, |E|\})$.

En el método de construcción, el espacio muestral está dado por duplas de la forma $(\tilde{X}, \tilde{\pi})$ donde la primera componente es un vector de estado \tilde{X} de G , y la segunda una permutación de los índices de los elementos de E , $\tilde{\pi} = (\tilde{\pi}_1, \dots, \tilde{\pi}_{|E|})$, que indica un ordenamiento de los enlaces. Se define además un índice k que corresponde al número de aristas operativas en un vector de estado: $\tilde{\pi}_1 = \dots = \tilde{\pi}_k = 1$ y $\tilde{\pi}_{k+1} = \dots = \tilde{\pi}_{|E|} = 0$.

La probabilidad de ocurrencia de la dupla $(\tilde{X}, \tilde{\pi})$ es

$$\rho(\tilde{X}, \tilde{\pi}) = \frac{\Pr(X = \tilde{X})}{k!(|E| - k)!}.$$

En la aplicación del método se obtiene una muestra de $\tilde{\pi}$ y se consideran todos los pares posibles $(\tilde{X}, \tilde{\pi})$ consistentes con la permutación. El vector de estado se elige de acuerdo con la probabilidad de ocurrencia de cada vector de estado de la red y la permutación se sortea considerando una variable Uniforme(0,1) en el espacio de permutaciones.

De acuerdo con el procedimiento de construcción (destrucción) la medida de confiabilidad $R_K(G)$ puede considerarse como una función del tiempo necesario para que la red cambie de un estado a otro.

La medida $R_K(G)$ se define como la probabilidad condicional de que el sistema opere dado $P_{\tilde{\pi}}$, y equivale al cociente entre la suma de la probabilidad de cada uno de los elementos en $P_{\tilde{\pi}}$ que cumplen $\Phi\{\tilde{x}\} = 1$ y la probabilidad de $P_{\tilde{\pi}}$:

$$R_K(G) = \frac{\sum_{k=0}^{|E|} \Phi(\tilde{X}^{(k)}) \rho(\tilde{X}^{(k)}, \tilde{\pi})}{\sum_{k=0}^{|E|} \rho(\tilde{X}^{(k)}, \tilde{\pi})} = \frac{\sum_{k=r}^{|E|} C_k^{|E|} \Pr(X = \tilde{X}^{(i)})}{\sum_{k=0}^{|E|} C_k^{|E|} \Pr(X = \tilde{X}^{(i)})} \quad (2.23)$$

ALGORITMO 6 Construcción secuencial

ENTRADA:

grafo G ,
 nodos objetivo K ,
 confiabilidad de los enlaces \vec{r} .

SALIDA:

estimación de la medida de confiabilidad: $\hat{R}_K(G)$.

- 1: Sortear $\tilde{\pi} = \{\tilde{\pi}_1, \dots, \tilde{\pi}_{|E|}\}$.
 - 2: **Para** $K = 0$ **hasta** $|E|$
 - 3: $\tilde{x}_{\pi_1} = \dots = \tilde{x}_{\pi_k} = 1$
 - 4: $\tilde{x}_{\pi_{k+1}} = \dots = \tilde{x}_{\pi_{|E|}} = 0$
 - 5: **Fin Para**
 - 6: Identificar el primer $r \in \{0, 1, \dots, |E|\}$ tal que $\Phi(\tilde{X}^r) = 1$
 - 7: Calcular $R_K(G)$
-

El Algoritmo 6 resume el método de Construcción. Los parámetros de entrada son la red $G = (V, E)$, el conjunto de nodos objetivo K y el vector \vec{r} de confiabilidad elemental de los enlaces, y retorna una estimación de la medida $R_K(G)$.

En la Línea 1 se sortea una permutación del vector $[1, 2, \dots, |E|]$ a partir de la cual se define un orden para cambiar el estado de los enlaces G .

En la Líneas 2 a 5 se realiza un proceso iterativo en el que se generan $|E|$ vectores de estado. El primero tiene un enlace operativo y $|E| - 1$ en estado de falla. El segundo vector tiene dos enlaces operativos y $|E| - 2$ en estado de falla, el k -ésimo tiene k enlaces operativos y $|E| - k$ en estado de falla.

En la Línea 6 se indica un procedimiento que implica identificar el valor $r \in [1, |E|]$ que cumple: $\Phi(x^{r-1}) = 0$ y $\Phi(x^r) = 1$.

En la Línea 7 se estima la medida de confiabilidad según la Ecuación 2.23.

Método de riesgo total

El método de riesgo total tiene aplicaciones variadas en distintos contextos como el área financiera o climática [34, 35]. Su aplicación para estimar la confiabilidad aparece en trabajos como [36, 37] de los años 1980 y 1992 respectivamente.

Como sugiere su nombre el método define una medida de riesgo asociada a sucesos que afectan negativamente el sistema en estudio. En particular en el modelo de confiabilidad en redes, un riesgo se identifica cuando todas las aristas de un cutset C_i fallan, y la medida de riesgo asociada, h_i , se define como la probabilidad de que todos los componentes de C_i fallen:

$$h_i = \prod_{l \in C_i} q_l. \quad (2.24)$$

Dado un grafo $G = (V, E)$, el método considera un cutset C_1 y simula el estado de las aristas que componen C_1 . Si en la primera simulación se generan todas las aristas de C_1 en estado de falla, la red falla y el proceso termina. Si en cambio al menos un enlace de C_1 opera, se fija el estado de las aristas de C_1 tal como se obtuvieron en la simulación anterior y se elige un nuevo conjunto de cutset C_2 . En cada iteración se calcula además el riesgo asociado a cada cutset, según Ecuación 2.24.

El proceso se repite hasta que todos los elementos de un cutset fallen o las aristas de G definan un sistema no aleatorio. Terminado el proceso iterativo se obtiene el estimador de la anticonfiabilidad como el riesgo total $H(G)$:

$$H(G) = \sum_{i=1}^r h_i,$$

donde r es el número de riesgos obtenidos o cutsets elegidos durante el procedimiento.

El Algoritmo 7 tiene como parámetros de entrada el grafo G , el vector de confiabilidad de los enlaces \vec{r} y la lista de cutsets de G y retorna una estimación de la medida de confiabilidad $R_K(G)$.

En la primera Línea del algoritmo se inicializa la variable H , que acumula la suma de los riesgos asociados a cada cutset en C .

En la Línea 2 se inicia el proceso eligiendo el primer cutset de G en función de algún criterio predefinido. Una opción es elegir el cutset de mayor riesgo.

En la Línea 3 se simula el estado de los enlaces que componen el cutset mediante sorteo de variables aleatorias de distribución Uniforme(0,1).

El bloque de código comprendido entre las Líneas 4 y 10 indica las iteraciones que continúan mientras el cutset elegido tiene al menos un enlace operativo. Se deja de iterar cuando la red falla.

En la Línea 5 se actualiza el estado de los enlaces del cutset en G , y en la Línea 6 se calcula el riesgo h asociado al cutset elegido. En la Línea 7 se acumula h en H .

En las Líneas 8 y 9 se elige un nuevo conjunto de cutset de G y se simula el estado de las aristas que lo conforman. Si el cutset no falla se continúa iterando.

ALGORITMO 7 Riesgo Total**ENTRADA:**

grafo G ,
 nodos objetivo K ,
 confiabilidad de los enlaces \vec{r} ,
 conjunto de cutset de G : C .

SALIDA:

estimación de la medida de confiabilidad: $\hat{R}_K(G)$.

- 1: $H = 0$
- 2: Elegir un cutset de C : C .
- 3: Simular estado de las aristas de C .
- 4: **Mientras** Opera alguna arista de C
- 5: Actualizo(C)
- 6: $h = \prod_{i \in C} q_i$
- 7: $H = H + h$
- 8: Elegir un cutset de C : C .
- 9: Simular estado de las aristas de C .
- 10: **Fin Mientras**
- 11: $R_K(G) = 1 - H$

Cuando culmina el proceso iterativo, el valor de H obtenido coincide con el riesgo total, y la medida de confiabilidad de la red es su complemento, $R_K(G) = 1 - H$ como indica la Línea 11 del Algoritmo 7.

La reducción de varianza que obtiene el método de riesgo total depende de la elección de los cutsets. Una alternativa es elegir en cada instancia el cutset de mayor riesgo pero esto requiere una búsqueda de costo computacional elevado, teniendo que revisar la lista de todos los cutsets de G , cuyo tamaño es de orden exponencial en el tamaño del grafo [29].

Muestreo de importancia

El muestreo de importancia provee una herramienta que permite reducir la varianza trasladando el problema de interés a un escenario en el que la probabilidad de ocurrencia del evento raro es de mayor magnitud.

El método propone realizar un cambio de medida con una nueva ley de probabilidad asociada, de forma que los eventos raros tiene alta probabilidad de ocurrencia bajo la nueva ley.

En el contexto de confiabilidad se tiene la medida $R_K(G) = E(\Phi(X))$ y una variable aleatoria X con función de densidad f_X . Si $X \in \mathbb{R}^n$ y las observaciones $X^{(i)}$ son i.i.d. de distribución f_X entonces el valor esperado de Φ es:

$$E(\Phi(X)) = \int \Phi(x)f(x)dx,$$

y la estimación es la media de las observaciones $\frac{1}{M} \sum_{i=1}^M \Phi(x^{(i)})$.

Se cumple además:

$$E(\Phi(X)) = \int \Phi(x) \frac{f(x)}{\tilde{f}(x)} \tilde{f}(x) dx = \int \Phi(x) L(x) \tilde{f}(x) dx, \quad (2.25)$$

y el estimador para la nueva expresión es:

$$\frac{1}{M} \sum_{i=1}^M \Phi(x^{(i)}) L(x^{(i)}), \quad (2.26)$$

donde $L(x) = \frac{f(x)}{\tilde{f}(x)}$ es el cociente de verosimilitud de la densidad $f(x)$ respecto de la densidad $\tilde{f}(x)$.

La simulación se genera en términos de la ley original utilizando $L(x)$ y el nuevo estimador definido en la Ecuación 2.26 es un estimador insesgado de $E(\Phi(X))$ y las observaciones $X^{(i)}$ son i.i.d. de distribución $\tilde{f}(x)$. Las varianzas asociadas a ambas densidades son:

$$\begin{aligned} \tilde{\sigma}^2(\Phi(X)L(X)) &= \tilde{E}[(\Phi(X)L(X))^2] - (E[\Phi(X)])^2 \\ \sigma^2(\Phi(X)) &= E[\Phi(X)^2] - (E[\Phi(X)])^2. \end{aligned}$$

El procedimiento no especifica cómo elegir la densidad para efectuar el cambio de medida que permita reducir la varianza, y, en efecto, ésta es la mayor dificultad que implica la aplicación del método. Sin embargo se define un tipo de estimador, *estimador de varianza cero*, que proporciona un valor de referencia para evaluar la calidad de un cambio de medida cualquiera.

Varianza cero

El **cambio de medida óptimo** [23] es aquél cuya probabilidad cumple:

$$L(X) = \frac{\partial P}{\partial \tilde{P}} = \frac{E[|\Phi(X)|]}{|\Phi(X)|}$$

y que equivale a

$$\tilde{f}(x) = \frac{f(x)|\Phi(x)|}{E[|\Phi(x)|]}.$$

En particular cuando $\Phi(x) \geq 0$ y se aplica el cambio de medida óptimo resulta que

$$\tilde{E}[(\Phi(X)L(X))^2] = (E[\Phi(X)])^2,$$

lo que genera que la varianza del estimador en \tilde{f}_x sea cero: $\tilde{\sigma}^2(\Phi(X)L(X)) = 0$.

El muestreo de importancia permite identificar un *estimador de varianza cero*, que se obtiene

al hacer un *cambio de medida de varianza cero*, pero esto implica conocer $E[|\Phi(X)|]$. Aunque no es posible determinarlo con exactitud es útil como valor de referencia.

Una noción para aproximarse al cambio “ideal” puede obtenerse considerando la definición del estimador: para que se mantenga acotado es necesario que $L(x)$ sea pequeño si $|\Phi(x)|$ es grande. Cuando $L(x) \leq \kappa, \forall x$ entonces se garantiza que el segundo momento se reduce por el factor ($\kappa < 1$): $\tilde{E}[(\Phi(X)L(X))^2] \leq \kappa E[(\Phi(X))^2]$. Si además $\Phi(X) = 1_A(X)$ entonces la varianza queda de la forma:

$$\tilde{\sigma}^2(1_A(X)\Phi(X)) = \tilde{E}[(1_A(X)L(X))^2] - Q^2 \leq \kappa^2 - Q^2,$$

por lo que κ tiene que ser parecido a Q .

En la práctica puede aproximarse el cambio de medida de varianza cero por medio de una aproximación a la función $R(\cdot)$ o restringir el estudio a una clase paramétrica y optimizar los parámetros. Existen aproximaciones por métodos adaptativos y no adaptativos para ambas posibilidades [23].

El cociente entre la mitad de la amplitud del intervalo de confianza y la cantidad Q a estimar es una medida de precisión que brinda una noción para identificar el cambio de medida a aplicar.

Se requiere que dicho cociente sea inferior a cierta tolerancia tol :

$$\frac{z_{\frac{\alpha}{2}} \frac{\tilde{\sigma}(\Phi(X)L(X))}{\sqrt{n}}}{Q} \leq tol$$

y para alcanzar la precisión deseada se necesita que $\frac{\tilde{\sigma}^2(\Phi(X)L(X))}{\sqrt{n}}$ sea proporcional a Q^2 , de forma que n sea proporcional a la varianza del estimador.

Método de cubrimiento

Este método propuesto en 1983 por Karp y Luby [38] es una variante de muestreo según importancia y muestreo estratificado y se basa en los mincuts de la red.

Sea C el conjunto de los mincuts del grafo $G = (V, E)$, w una función de ponderación no negativa de los mincuts de G , y un espacio universal (U, w) que contiene los eventos de falla F del sistema. Los elementos del espacio universal son pares de la forma (X, C) , con X un vector de estado de $G(V, E)$ y C un mincut. Nótese que la definición de los pares (X, C) no establece una relación inyectiva, pues distintos vectores de estado pueden asociarse a un mismo mincut.

Para que exista relación única entre los estados de falla del conjunto F y los mincuts en U se define una relación que permite asignar a cada mincut un único X . Para cada nodo $s \in K$ se define un conjunto A_s como el conjunto de todos los nodos de G alcanzables desde s por caminos operativos. Se define el mincut $C(X)$ como el conjunto conformado por los enlaces que conectan los nodos de A_s con los de $V - A_s$. Esta relación asegura que cada elemento en

F tenga un único correspondiente en U y es verificable en tiempo polinomial [29].

Definida la relación entre F y U , el método de cubrimiento permite estimar la medida de anticonfiabilidad a partir de una muestra en U . La estimación se obtiene multiplicando los elementos de U que están incluidos en F por $w(U)$, lo que asegura que es un estimador insesgado de $Q_K(G)$ [23].

La función de ponderación cumple:

- $w(F) = \Pr(F) = Q_K(G)$
- Dado un mincut C , se define $w(C)$, un valor de ponderación total de los elementos de U con segunda componente igual a C . Se calcula como $w(C) = \prod_{e \in C} q_e$.
- $w(U) = \sum_{C \in \mathcal{C}} w(C)$ se calcula en tiempo polinomial.
- Se puede decidir en forma eficiente si un elemento de U pertenece a F .
- Se define una cota $M: \frac{w(U)}{w(F)} \leq M$

ALGORITMO 8 Método de cubrimiento

ENTRADA:

grafo G ,

nodos objetivo K ,

confiabilidad de los enlaces \vec{r} ,

conjunto de cutset de G : C .

SALIDA:

estimación de la medida de anticonfiabilidad: $\hat{Q}_K(G)$.

- 1: **Para** $C = 1$ **hasta** $|C|$
 - 2: $w(C) = \prod_{e \in C} q_e$
 - 3: **Fin Para**
 - 4: $w(U) = \sum_{C \in \mathcal{C}} w(C)$
 - 5: Sortear un mincut C con probabilidad $\frac{w(C)}{w(U)}$.
 - 6: $\forall e \in C \tilde{x}_e = 0$
 - 7: $\forall e \notin C$ sorteo $\tilde{x}_e = 1$ con probabilidad r_e
 - 8: **Si** $C = C(\tilde{x})$ **entonces**
 - 9: $Q_K(G) = w(U)$
 - 10: **Si no**
 - 11: $Q_K(G) = 0$
 - 12: **Fin Si**
-

La implementación del método de cubrimiento se resume en el Algoritmo 8. El algoritmo retorna una estimación de la medida de anticonfiabilidad $Q_K(G)$. Los parámetros de entrada son el grafo $G = (V, E)$, el conjunto de nodos objetivo K , el vector de confiabilidad de los enlaces \vec{r} y el conjunto de mincuts de la red C

En las Líneas 1 a 3, se define el valor de la función de ponderación para cada mincut del grafo. En la Línea 4 se define un ponderador de U como la suma de los ponderadores de cada mincut y en la Línea 5 se sortea un mincut C con probabilidad proporcional al peso del mincut respecto al total de mincuts del grafo: $\frac{w(C)}{w(U)}$. A mayor probabilidad de falla del mincut mayor es la probabilidad de ser elegido.

En las Líneas 6 a 8 se define el vector de estados asociado a C , de forma que el estado de los enlaces que integran el mincut fallan y el resto operan.

Si el mincut sorteado coincide con un conjunto separador asociado a \tilde{x} entonces el sistema falla y se define la medida de anticonfiabilidad: $Q_K(G) = w(U)$. Si en cambio el conjunto sorteado no es separador de los nodos objetivo entonces el sistema tiene estado operativo, $Q_K(G) = 0$ como indica la Línea 11.

El método de cubrimiento permite reducir la varianza respecto de CMC , pero su aplicación es costosa dado que requiere enumerar e identificar la lista de mincuts de G . El cálculo y almacenamiento de esta información es compleja y acota su aplicación a algunas topologías [29].

Método de reducción recursiva de la varianza.

El método *RVR* fue presentado en 1994 por Cancela y El Khadiri [39].

A partir de un grafo $G = (V, E)$ con espacio de vectores de estado $\Omega_G = \{0, 1\}^m$ con m el número de enlaces. El procedimiento aplica una partición P_G de Ω_G , en función de la cual define la probabilidad de funcionamiento del sistema.

Se define la partición P_G de Ω_G de acuerdo con el estado de los enlaces de la red:

$$P_G = \{A_O, A_F\} \cup \{A_w; w \in W\},$$

y se cumple:

1. Si $A_i \neq A_j$ entonces $A_i \cap A_j = \emptyset, \forall A_i, A_j \in P_G$.
2. $\bigcup_{A_i \in P_G} A_i = \Omega_G$.
3. Si $A_O \neq \emptyset$, entonces A_O está conformado solamente por estados operativos del sistema:

$$R_K(G_O) = 1, \forall G_O \in A_O.$$

4. Si $A_F \neq \emptyset$, entonces A_F está formado sólo por estados de falla del sistema:

$$R_K(G_F) = 0, \forall G_F \in A_F.$$

5. Para cada $w \in W$ existe $G_w = (V, E_w) \in A_w$ tal que $|E_w| < |E|$ y $R_K(G_w)$ en el intervalo $(0, 1)$.

Los puntos 1 y 2 indican propiedades generales de partición de conjuntos. Los puntos 3 y 4 indican estados donde la red es un sistema determinista, en el sentido de que el estado

del sistema es fijo y todos los enlaces aleatorios son irrelevantes. Por último en el punto 5 se define el conjunto de estados aleatorios de la red, cuyo estado está sujeto a la aleatoriedad de los componentes.

Dados los componentes de P_G se define, en términos generales, la probabilidad de que un vector de estado X de G pertenezca a un componente de la partición:

1. $P_O = \Pr(X \in A_O)$, donde P_O es la probabilidad de funcionamiento incondicional
2. $P_F = \Pr(X \in A_F)$, donde P_F es la probabilidad de falla incondicional
3. Se define una variable aleatoria W , tal que $\Pr(W = w) = \Pr(X \in A_w | X \in \Omega_G \setminus (A_O \cup A_F))$, es la probabilidad de ocurrencia de un estado perteneciente al conjunto $A_w \subseteq \Omega_G \setminus (A_O \cup A_F)$.

Los componentes de P_G pueden definirse en función de los minpaths o mincuts del grafo G . Si P es el conjunto de los minpaths de G y P_i es un minpath arbitrario en P entonces definimos A_O , A_w y A_F de la siguiente forma:

- Si P_i es un minpath de G , A_O es el conjunto de vectores de estado de la red tal que “todas las aristas de P_i operan”,
- A_w es el conjunto de vectores de estado de la red tal que “la w -ésima arista del minpath P_i falla y las $w - 1$ aristas anteriores operan”,
- $A_F = \emptyset$,

y la probabilidad de que un estado X pertenezca a un componente de P_G definido en términos de un minpath cumple:

1. $P_O = \Pr(X \in A_O) = \prod_{j \in P_i} r_j$, con r_j la confiabilidad elemental de la j -ésima arista de P_i .
2. Se define una variable aleatoria $\Pr(W = w) = \prod_{k=1}^{w-1} r_k (1 - r_w)$, donde las aristas $1, 2, \dots, w - 1$ de P_i están operativas y la w -ésima falla.

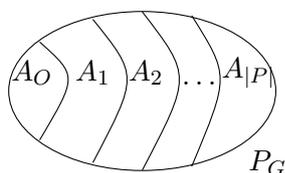


Figura 2.11: Partición de Ω_G basada en paths

Ejemplo 2.5 En el ejemplo de la Figura 2.12 se considera un grafo $G = (V, E)$, $|V| = 7$ y $|E| = 11$ con nodos objetivo $K = \{s, t\}$. Un minpath de G se denota en la figura con una

línea de mayor grosor y corresponde al conjunto de enlaces, $\{e_1, e_2, e_3\}$, que conectan los nodos objetivo $\{s, t\}$.

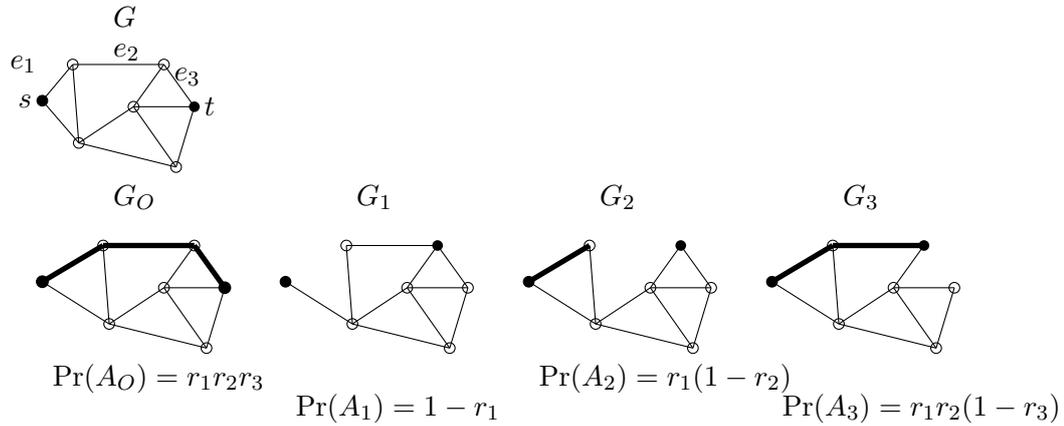


Figura 2.12: Partición basada en minpaths

La partición consta de 4 componentes:

- A_0 : El conjunto de subgrafos de G donde e_1, e_2, e_3 están operativas. La probabilidad de que una configuración de G pertenezca a A_0 es $\Pr(X \in A_0) = \prod_{i=1}^3 r_i$.
- A_1 : El conjunto de subgrafos de G donde e_1 falla. La probabilidad de que una configuración de G pertenezca a A_1 es $\Pr(X \in A_1) = 1 - r_1$.
- A_2 : El conjunto de subgrafos de G donde e_2 falla y e_1 está operativa. La probabilidad de que una configuración de G pertenezca a A_2 es $\Pr(X \in A_2) = r_1(1 - r_2)$.
- A_3 : El conjunto de subgrafos de G donde e_3 falla y e_1, e_2 están operativas. La probabilidad de que una configuración de G pertenezca a A_3 es $\Pr(X \in A_3) = r_1 r_2(1 - r_3)$.

Por otro lado, la partición puede definirse en términos de mincuts de la red. Si C es el conjunto de todos los mincuts de G , $C = \{C_1, C_2, \dots, C_{|C|}\}$ la partición P_G se define a partir de $C_i \in C$ arbitrario:

- A_F es el conjunto de los subgrafos donde “todas las aristas de C_i fallan”,
- A_w es el conjunto de vectores de estado de G tal que “la w -ésima arista del cutset C_i opera y las primeras $w - 1$ aristas de C_i fallan”,
- $A_O = \emptyset$,

La probabilidad de que un vector de estado X pertenezca a una componente de la partición se asocia a la probabilidad de falla de los enlaces de C_i :

1. $P_F = \Pr(X \in A_F) = \prod_{j \in C_i} (1 - r_j)$, con r_j la confiabilidad elemental de la j -ésima arista de C_i .
2. $\Pr(W = w) = \prod_{k=1}^{w-1} (1 - r_k)(r_w)$, donde las aristas $1, 2, \dots, w - 1$ de C_i fallan y la w -ésima opera.

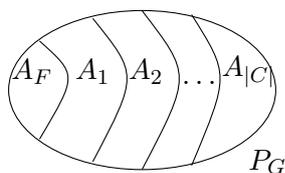


Figura 2.13: Partición de Ω_G basada en mincut

Ejemplo 2.6 Sea G el grafo de la Figura 2.14, con nodos objetivo $K = \{s, t\}$ y un mincut $C = \{e_1, e_2\}$. Los enlaces operativos se indican en la figura con una línea más densa, los enlaces en falla se eliminan. La partición que se obtiene a partir de C consta de 3 componentes $P_G = \{A_F, A_1, A_2\}$ donde:

- A_F es la partición que contiene todos los vectores de estado de G tal que fallan todos los componentes del cutset. La probabilidad de que un vector de estado de G pertenezca a A_F es

$$\Pr(X \in A_F) = \prod_{i=1}^2 q_i.$$

- A_1 es el conjunto de todos los vectores de estado de G tal que la primera arista de C opera. La probabilidad de que un vector de estado de G pertenezca a A_1 es

$$\Pr(X \in A_1) = r_1.$$

- A_2 es el conjunto de todos los vectores de estado de G tal que la primera componente de C falla y la segunda opera. La probabilidad de que un vector de estado de G pertenezca a A_2 es

$$\Pr(X \in A_2) = (1 - r_1)r_2.$$

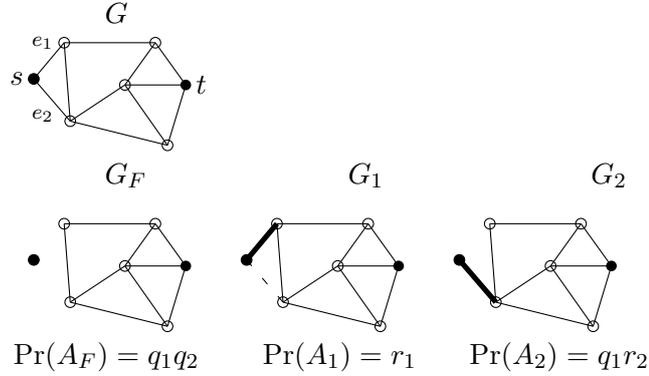


Figura 2.14: Partición basada en cortes

De aquí en adelante cuando se haga referencia a P_G de Ω_G se considerará la partición definida en función de un mincut, salvo que se explicita lo contrario.

Una vez que se identifica una partición de Ω_G y se calcula la probabilidad de que un vector de estado de G pertenezca a uno u otro componente de P_G , se define una fórmula recursiva para el cálculo de la medida de anticonfiabilidad $Q_K(G)$.

$$Q_K(G) = Q_{C(G,K)} + (1 - Q_{C(G,K)}) \sum_{i=1}^{|C(G,K)|} \frac{p(A_i)}{(1 - Q_{C(G,K)})} Q_{K_i}(G_i) \quad (2.27)$$

donde:

- $Q_K(G)$ es la medida de anticonfiabilidad de la red.
- $C(G, K) = \{e_1, e_2, \dots, e_{C(G,K)}\}$ es el cutset de G .
- $Q_{C(G,K)}$ es la probabilidad de falla del cutset $C(G, K)$.
- A_i es el evento “fallan las primeras $i - 1$ aristas de $C(G, K)$ y opera la i -ésima”.
- $G_i(V_i, E_i)$ y K_i son el grafo y el conjunto de nodos objetivo que se obtiene de $G(V, E)$ y K , condicionados al evento A_i .

De la expresión 2.27 se deriva un operador recursivo para estimar la medida de anticonfiabilidad:

$$F(G, K) = Q_{C(G,K)} + (1 - Q_{C(G,K)}) \sum_{i=1}^{|C(G,K)|} 1_{\{U \in J_i\}} F(G_i, K_i) \quad (2.28)$$

donde:

- U es una variable aleatoria con distribución Uniforme(0,1).

- J_i con $1 \leq i \leq |C(G, K)|$ es una secuencia de intervalos disjuntos de amplitud $\frac{\Pr(A_i)}{1 - Q_{C(G, K)}}$ tal que $\bigcup_{i=1}^{|C(G, K)|} J_i = [0, 1]$
- $p(A_i) = \Pr(A_i) = \prod_{k=1}^{i-1} (1 - r_k)(r_i)$.

Nótese que en cada sumando $\sum_{i=1}^{|C(G, K)|} 1_{\{U \in J_i\}} F(G_i, K_i)$ de la Ecuación 2.28 sólo una de las $|C(G, K)|$ indicatrices vale 1 de forma que en cada paso recursivo se considera sólo uno de los componentes de la partición. El subgrafo G_i es tal que fallan los primeros $i - 1$ enlaces del cutset, y el i -ésimo enlace está operativo.

El método *RVR* se basa en la Ecuación 2.28. En analogía con *CMC*, *RVR* sorteja una muestra de configuraciones de G independientes, pero en lugar de realizar sorteos independientes para determinar el estado de los enlaces genera estados condicionados a la falla de enlaces del cutset.

El estimador de la medida de anticonfiabilidad se define como la media de las observaciones de $F(G, K)$ [39]:

$$\hat{Q}_K(G) = \frac{\sum_{i=1}^N F^i(G, K)}{N}; \quad (2.29)$$

y el estimador insesgado de la varianza, $Var(\hat{Q}_K(G)) = \frac{Var(F(G, K))}{N}$ [39] como:

$$\hat{V} = \frac{\sum_{i=1}^N (F^i(G, K) - \hat{Q}_K(G))^2}{N(N-1)}. \quad (2.30)$$

ALGORITMO 9 *RVR*

ENTRADA:

grafo G ,
 nodos objetivo K ,
 confiabilidad de los enlaces \vec{r} ,
 número de iteraciones N .

SALIDA:

estimación de la medida de anticonfiabilidad: $\hat{Q}_K(G)$,
 estimación de la varianza: \hat{V}_Q .

- 1: $Y = []$
 - 2: **Para** $i = 1$ **hasta** N
 - 3: $Y_i = F(G, K, r)$
 - 4: **Fin Para**
 - 5: $\hat{Q}_K(G) = \frac{\sum_{i=1}^N Y_i}{N}$
 - 6: $\hat{V}_Q = \frac{\sum_{i=1}^N (Y_i - \hat{Q}_K(G))^2}{N(N-1)}$
-

En el Algoritmo 9 se resumen los pasos del método *RVR*. Los parámetros de entrada son el grafo $G = (V, E)$, el conjunto de nodos objetivo K , la confiabilidad elemental de los enlaces,

\vec{r} y el tamaño de muestra N , y los parámetros de salida son la estimación de $Q_K(G)$ y su varianza.

En la Línea 1 se define Y , un vector en \mathbb{R}^N en donde se almacenan N valores simulados de $F(G, K)$. Y_i es el i -ésimo componente de Y que contiene el valor obtenido en la i -ésima iteración.

En la Línea 2 se invoca el Algoritmo 10 mediante el cual se simula un estado de la red en base al operador recursivo 2.28.

Por último, finalizadas las iteraciones se estiman los parámetros. La estimación de la medida de anticonfiabilidad $\hat{Q}_K(G)$ y la estimación de su varianza \hat{V}_Q se calculan según las Ecuaciones 2.29 y 2.30.

ALGORITMO 10 $F = F(G, K, \vec{r})$

ENTRADA:

grafo G ,
 nodos objetivo K ,
 confiabilidad de los enlaces \vec{r} .

SALIDA:

simulación de un estado de G : F .

- 1: **Si** $|K| = 1$ **entonces**
 - 2: $F = 0$
 - 3: **Fin Si**
 - 4: **Si** G falla **entonces**
 - 5: $F = 1$
 - 6: **Si no**
 - 7: Obtener $C = \{e_1, e_2, \dots, e_{|C|}\}$
 - 8: $Q_C = \prod_{i=1}^{|C|} q_i$
 - 9: Sortear U Uniforme(0,1)
 - 10: Elegir J_i que cumple $U \in J_i$
 - 11: Construir $G_i = (G - e_1 - e_2 - \dots - e_{i-1}) * e_i$
 - 12: Construir K_i
 - 13: $F = Q_C + (1 - Q_C)RVR(G_i(V, E), K_i, \vec{r})$
 - 14: **Fin Si**
-

En el Algoritmo 10 aplica el operador recursivo $F(G, K)$. A partir de la red $G = (V, E)$ el conjunto de nodos objetivo K y la confiabilidad de las aristas \vec{r} retorna una realización del operador F .

En la Línea 1 se evalúa el caso en que el conjunto objetivo tiene un único elemento. Si es así el sistema siempre está operativo y la medida de anticonfiabilidad es 0.

En la Línea 4 se evalúa si el sistema está en estado de falla. En este caso la anticonfiabilidad vale 1.

Si la red no es un sistema determinista se aplica una búsqueda recursiva. En la Línea 7 se identifica un cutset C de la red.

En la Línea 8 se calcula la probabilidad de falla del cutset definida como el producto de la

anticonfiabilidad elemental de sus aristas.

En la Línea 9 se sortea una variable aleatoria U Uniforme(0,1) en función de la cual se determina en qué componente de la partición continúa el proceso.

En la Línea 10 se identifica el intervalo J_i tal que $U \in J_i$. El intervalo es de amplitud proporcional a la probabilidad de ocurrencia del componente i -ésimo de P_G .

En la Línea 11 se construye G_i , resultando de eliminar las primeras $i - 1$ aristas de C y fijar la i -ésima arista operativa, o de manera equivalente contraer el enlace.

En la Línea 12 se construye K_i . Eventualmente K_i coincide con K o bien tiene menos elementos, como resultado de aplicar reducciones a la red [40].

Por último en la Línea 13 se realiza la llamada recursiva considerando la Ecuación 2.28.

El proceso recursivo continúa hasta que la subred G_i generada es determinista (siempre está en estado de falla u operativo).

La precisión de la estimación y el desempeño computacional depende de la elección de los cutsets o pathsets [29].

La recursión tiene un número finito de pasos [39] y su desempeño es bueno sólo en algunas clases de grafos y niveles de confiabilidad elemental. El proceso recursivo involucra uso excesivo de memoria y en ocasiones el tiempo de cálculo también es excesivo. El bloque recursivo del algoritmo tiene una complejidad de almacenamiento de $O(m(m+n))$ y el tiempo computacional en el peor de los casos es de orden $O((mk)^2n)$, con $|E| = m$, $|V| = n$, $|K| = k$ [23].

La aceleración del método respecto a *CMC* puede mejorarse aplicando reducciones serie-paralelo que preservan la medida de confiabilidad del sistema simplificando la topología [14].

Capítulo 3

Confiabilidad diámetro acotada

3.1. Introducción

El crecimiento acelerado de las redes de comunicación, su mayor complejidad y la variedad de servicios que ofrece a sus usuarios han abierto nuevas y diversas líneas de investigación. Los usuarios, en tanto demandan un servicio y enfrentan un costo para acceder a él, exigen estándares de calidad (QoS) y como contrapartida, para satisfacer esa demanda la red debe ofrecer comunicación eficiente y eficaz.

Existen aplicaciones en donde además de eficacia en la comunicación se requieren retardos mínimos en la transmisión, apelando a la eficiencia de la red. Un ejemplo es la comunicación por videoconferencia, donde es necesario que el delay existente entre el emisor del mensaje y el receptor sea menor que un valor predeterminado. Si la tolerancia para el retardo es T y el tiempo de transmisión por enlace es t entonces la comunicación debe recorrer a lo sumo D enlaces tal que $tD \leq T$. Otra aplicación que considera los retardos se asocia a la optimización de transmisión de paquetes en redes reales peer-to-peer (*P2P*) [41, 42].

Este nuevo requisito en el servicio se resume en una nueva medida de confiabilidad en redes que considera una cota superior para la distancia entre los nodos que se desea comunicar, es la medida de confiabilidad diámetro acotada, denotada como $R_{\{K,D\}}(G)$. Esta medida, definida por Petingi y Rodríguez [43] en el año 2001 generaliza el concepto de medida de confiabilidad clásica $R_K(G)$, estableciendo que una red está operativa si los nodos objetivo se comunican por caminos cuyo largo no supera la cota.

En este capítulo se presenta la teoría general asociada a la medida de confiabilidad diámetro acotada, extendiendo los conceptos definidos en el Capítulo 2. Vale destacar que el desarrollo teórico respecto de la medida diámetro acotada es incipiente si se lo compara con la medida clásica.

En la Sección 3.2 se presenta la definición de confiabilidad diámetro acotada y detalla el modelo estocástico. En la Sección 3.3 se consideran métodos de cálculo exacto, con especial énfasis en algoritmos de factorización.

En la Sección 3.4 se introducen métodos aproximados desarrollando en particular una variante

del método *RVR*.

Para finalizar, en la Sección 3.5 se resumen antecedentes que investigan características de la medida $R_{K,D}(G)$. En la Sección 3.6 se detallan problemas abiertos.

3.2. Modelo probabilístico

El modelo probabilístico para la medida de confiabilidad diámetro acotada es una generalización del modelo clásico. Dada una red $G = (V, E)$ con enlaces no dirigidos se asumen nodos perfectos y comportamiento de falla independiente con probabilidad conocida en los enlaces. Se identifica el conjunto de nodos objetivo K , $K \subseteq V$, de los que se requiere conectividad y se incorpora, respecto del modelo clásico, una restricción de diámetro. La conexión entre nodos objetivo debe ocurrir a través de caminos operativos cuya longitud es a lo sumo D , donde $D \in \mathbb{Z}^+$ y $D \in [1, n - 1]$ con $n = |V|$ (la longitud coincide con el número de aristas que componen el camino).

La medida de confiabilidad, *confiabilidad K -terminal de G con restricción de diámetro D* , se denota por $R_{K,D}(G)$ y se define como la probabilidad de que los nodos objetivo en K estén comunicados por caminos de largo no mayor que D una vez eliminados los enlaces en estado de falla.

De acuerdo con la conformación del conjunto objetivo K se definen además los casos particulares, confiabilidad diámetro acotada V -terminal, $R_{V,D}(G)$ y la confiabilidad diámetro acotada s - t -terminal, $R_{\{s,t\},D}(G)$.

Nótese que la medida de confiabilidad clásica $R_K(G)$ es un caso particular de $R_{K,D}(G)$ con $D \geq n - 1$. De esta forma todo camino operativo que conecte los nodos objetivo es un estado operativo del grafo sin importar su longitud. Cuando se impone la restricción de diámetro, el espacio de estados operativos se reduce a un subconjunto del espacio de estados global de G .

Ejemplo 3.1 En los grafos de la Figura 3.1, se considera $D = 4$ y $K = \{a, b, c, d\}$.

En el grafo de la izquierda todos los enlaces están operativos y la comunicación entre los nodos de K es exitosa porque existen caminos que comunican los nodos objetivo de largo 4 y menos.

En cambio, en el grafo de la derecha no existe un camino entre los nodos objetivo d y b que cumpla con la restricción de diámetro. El camino más corto que los comunica consta de 6 enlaces y por lo tanto la red falla.

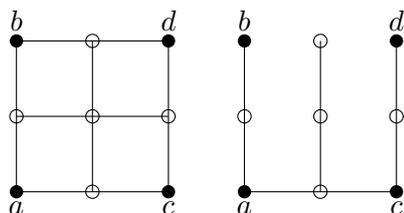


Figura 3.1: Ejemplo Grilla 3×3

Ejemplo 3.2 Sea $G = (V, E)$ el grafo de la Figura 3.2 con nodos objetivo $\{s, t\}$. Se denota el enlace e en función de sus extremos tal que $e = (x, y)$ es el enlace con nodos extremos x e y , y un camino como la secuencia de enlaces que lo componen.

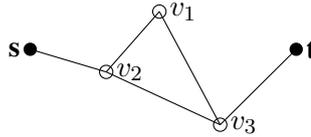


Figura 3.2: Confiabilidad s - t terminal diámetro acotada: caminos operativos

El conjunto de caminos operativos del grafo G depende de la cota de diámetro que se define. En este caso D puede tomar los valores $\{1, 2, 3, 4\}$. (Los casos $D > 4$ quedan contemplados cuando se asume $D = 4$, donde la medida coincide con la medida de arista confiabilidad clásica).

De acuerdo con el grafo de la Figura 3.2, existen dos caminos, p_1, p_2 , que conectan los nodos objetivo $K = \{s, t\}$; $p_1 = \{(s, v_1), (v_1, v_3), (v_3, t)\}$ y $p_2 = \{(s, v_1), (v_1, v_2), (v_2, v_3), (v_3, t), t\}$. Si se considera $D = 4$ el grafo opera si al menos uno de los caminos opera, pues ambos caminos son factibles.

Si $D = 3$ sólo es factible el camino p_1 , pues p_2 consta de 4 enlaces.

Si $D \leq 2$ el grafo nunca opera, pues no existen caminos que conecten s y t de longitud inferior a 3.

Retomando el modelo, el estado de cada enlace de la red se representa con una variable aleatoria Bernoulli, donde la probabilidad de que $e \in E$ opere es r_e y la probabilidad de que falle es $q_e = 1 - r_e$. El estado global de la red se describe mediante un vector de configuración X , a los que se asocia una función de estructura $\Phi(X)$.

La nueva medida de confiabilidad $R_{K,D}(G)$ modifica la definición de estado operativo y de conjunto corte de G incorporando los conceptos D -conectado, D -pathset y D -minpath.

Definición 3.2.1 Dado $G = (V, E)$, el conjunto de nodos objetivo K y una cota de diámetro D , dos nodos de K están **D -conectados** si existe entre ambos al menos un camino operativo de largo menor o igual a D .

Si dado un vector de estado x , todos los nodos objetivo están D -conectados entonces la red está operativa y $\Phi(x) = 1$. En éste caso los enlaces operativos en x definen un D -pathset.

Definición 3.2.2 Dado $G = (V, E)$, el conjunto de nodos objetivo K y una cota de diámetro D , un **D -pathset** es un estado operativo de G si recubre un subgrafo con K -diámetro menor o igual a D donde el K -diámetro de G es la distancia máxima entre los nodos de K [41].

Si P es un D -pathset entonces $x_j = 1, \forall x_j \in P$ y $\Phi(x) = 1$

Definición 3.2.3 Dado $G = (V, E)$, el conjunto de nodos objetivo K y una cota de diámetro D , un **D -minpath** P genera un estado operativo. Si x es tal que $x_j = 1, \forall x_j \in P$ y $x_i = 0, \forall x_i \notin P$, entonces $\Phi(x) = 1$ y $\Phi((x, 0_j)) = 0$

Por otro lado, la nueva caracterización mantiene el concepto de conjunto de corte. Sin embargo, los cortes posibles de la red varían como también cambia el procedimiento en que los conjuntos de corte deben ser identificados.

Definición 3.2.4 Un *D-cutset* C es un conjunto de enlaces cuya eliminación asegura un estado de falla de la red. Si x es un vector tal $x_j = 0, \forall x_j \in C$ entonces $\Phi(x) = 0$.

Definición 3.2.5 Un *D-minicut* C es un D -cutset minimal en el sentido que si al menos un elemento de C no se elimina de la red ésta permanece operativa. Si x es un estado tal que $x_j = 0, x_j \in C$ y $x_k = 1, x_k \notin C$, entonces $\Phi(x) = 0$ y $\Phi((x, 1_j)) = 1$.

Dada una cota de diámetro D , un D -cutset admite una distinción de acuerdo con la forma en que interrumpe la comunicación. La eliminación de un cutset separa la red en dos o más componentes conexas interrumpiendo la conexión de los nodos objetivo, o bien sin separar la red, evita que la comunicación ocurra por D -pathsets.

Ejemplo 3.3 Dado el grafo de la Figura 3.3, en un modelo sin cotas de diámetro el conjunto de cutsets es $C = \{C_1, C_2, C_3, C_4\}$, con $C_1 = \{(s, v_2)\}$, $C_2 = \{(v_2, v_3), (v_2, v_1)\}$, $C_3 = \{(v_3, t)\}$, $C_4 = \{(v_2, v_3), (v_1, v_3)\}$.

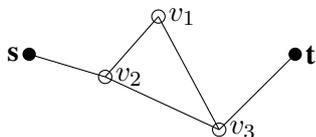


Figura 3.3: Confiabilidad s - t terminal diámetro acotada: conjuntos de corte

En cambio si $D = 2$ la red nunca opera y $C = \emptyset$ en el sentido que no hay estado operativo que interrumpir.

En cambio, en el grafo puente de la Figura 3.4 para $D = 2$, los D -cutsets posibles son: $C_1 = \{(s, u), (s, v)\}$, $C_2 = \{(t, u), (t, v)\}$, $C_3 = \{(s, v), (u, t)\}$ y $C_4 = \{(s, u), (v, t)\}$. En particular, los dos últimos D -cutsets no corresponden a conjuntos separadores pero impiden que los nodos objetivo se comuniquen por caminos de largo 2.

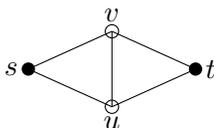


Figura 3.4: Grafo puente

Identificados los conjuntos de D -minpaths o D -mincuts la medida $R_{K,D}(G)$ queda definida en función de la probabilidad de funcionamiento o falla respectivamente. Dada $G = (V, E)$, el conjunto de nodos objetivo K y la cota de diámetro D :

- Sea P^D el conjunto de todos los D -minpaths de G , $P^D = \{P_1^D, P_2^D, \dots, P_{|P^D|}^D\}$
- Sea P_i^D un D -minpath, $i = 1, 2, \dots, |P^D|$, tal que $P_i^{D,O}$ es el evento “todas las aristas de P_i^D están operativas”.
- $\Pr(P_i^{D,O}) = \prod_{j \in P_i^D} r_j$, con r_j la confiabilidad elemental de la j -ésima arista de P_i^D .
- Si al menos ocurre uno de los eventos $P_i^{O,D}$ entonces el sistema está operativo y se cumple

$$R_{K,D}(G) = Prob \left(\bigcup_{P_i^D \in P} P_i^{O,D} \right)$$

Por analogía se define la medida de anticonfiabilidad diámetro acotada $Q_{K,D}(G)$ en función de un conjunto de D -cutssets:

- Sea C^D el conjunto de todos los mincuts de G , $C^D = \{C_1^D, C_2^D, \dots, C_{|C^D|}^D\}$
- Sea C_i^D un mincut, $i = 1, 2, \dots, |C^D|$, tal que $C_i^{D,F}$ es el evento “todas las aristas de C_i^D fallan”
- $\Pr(C_i^{D,F}) = \prod_{j \in C_i^D} (1 - r_j)$, con r_j la confiabilidad elemental de la j -ésima arista de C_i^D .
- Si al menos ocurre uno de los eventos $C_i^{D,F}$ entonces el sistema falla:

$$R_{K,D}(G) = 1 - Prob \left(\bigcup_{C_i^D \in C^D} C_i^{D,F} \right).$$

3.3. Métodos de cálculo exacto

En el caso general el cálculo exacto de $R_{K,D}(G)$ es complejo y el costo computacional asociado a su implementación es computacionalmente prohibitivo. Sin embargo existen casos triviales donde el cálculo es inmediato [44].

Para $D = 1$ el problema es sencillo: si el subgrafo de G inducido por los nodos de K no es un clique (grafo completo) entonces $R_{K,D}(G) = 0$, pues existen nodos no conectados o conectados por caminos de largo mayor a 1. Trivialmente, si el grafo inducido es un clique la confiabilidad equivale a:

$$R_{K,D}(G) = \prod_{k_i, k_j \in K} r_{k_i, k_j}.$$

Cuando $D = 2$ y $K = \{s, t\}$ la medida de confiabilidad puede calcularse en tiempo lineal en el tamaño de G [44]. Los únicos caminos relevantes entre los nodos objetivo son:

$$P_{s,t} = \{s, t\} \cup \{\{s, v, t\}, v \in V \setminus \{s, t\}, (s, v), (v, t) \in E\}.$$

Se denota $P_{s,t} = \{P_1, P_2, \dots, P_{|P_{s,t}|}\}$. Se define el evento P_i^O como “todos los enlaces de P_i operan” y \bar{P}_i^O su complemento. La medida de confiabilidad se deriva aplicando la técnica “uniones disjuntos” [3, 45].

$$R_{\{s,t\},2}(G) = Prob \left(\bigcup_{P_i \in P_{s,t}} P_i^O \right) = Prob \left(\bigcup_{P_i \in P_{s,t}} \left(P_i^O \cap \left(\bigcap_{j < i} \bar{P}_j^O \right) \right) \right). \quad (3.1)$$

Nótese que el cambio en la formulación permite trasladar una expresión de eventos no disjuntos,

$$Prob \left(\bigcup_{P_i \in P_{s,t}} P_i^O \right), \quad (3.2)$$

a una expresión de eventos disjuntos,

$$Prob \left(\bigcup_{P_i \in P_{s,t}} \left(P_i^O \cap \left(\bigcap_{j < i} \bar{P}_j^O \right) \right) \right). \quad (3.3)$$

Retomando la definición de $R_{\{s,t\},2}(G)$, la expresión final en 3.1, unión de eventos disjuntos, permite expresar la probabilidad de la unión como suma de probabilidades:

$$R_{\{s,t\},2}(G) = \sum_{P_i \in P_{s,t}} Prob \left(P_i^O \cap \left(\bigcap_{j < i} \bar{P}_j^O \right) \right).$$

Dada la independencia del estado de los enlaces y de los D-minpaths la probabilidad de la intersección del estado de los enlaces es el producto de las probabilidades:

$$R_{\{s,t\},2}(G) = \sum_{P_i \in P_{s,t}} Prob(P_i^O) \left(\prod_{j < i} Pr(\bar{P}_j^O) \right).$$

Cada evento P_i^O es la intersección de eventos independientes y la probabilidad de P_i^O es el producto de la confiabilidad elemental de los enlaces que lo componen. Teniendo en cuenta que el conjunto $P_{s,t}$ tiene un D-minpath formado por un único enlace, $e = (s, t)$, y los demás elementos son caminos de largo 2, la medida de confiabilidad es $R_{\{s,t\},2}(G)$:

$$R_{\{s,t\},2}(G) = r_{s,t} + (1 - r_{s,t}) \sum_{i \in V \setminus \{s,t\}} r_{s,i} r_{i,t} \prod_{j < i, j \in V \setminus \{s,t\}} (1 - r_{s,j} r_{j,t}). \quad (3.4)$$

La medida s - t -terminal también puede interpretarse por su complemento. En la Ecuación 3.5 se formula la medida $R_{\{s,t\},2}(G)$ en base a la probabilidad de falla de los caminos que conectan s y t .

$$R_{\{s,t\},2}(G) = \begin{cases} 1 - R', & (s, t) \notin E, \\ 1 - (1 - r_{s,t})R', & \text{si } (s, t) \in E, \end{cases} \quad (3.5)$$

donde

- $R' = \prod_{i=1}^l (1 - r_{(s,u_i)}r_{(u_i,t)})$
- $r_{(s,u_i)}r_{(u_i,t)}$ es la probabilidad de que el camino de 2 aristas $\{s, u_i, t\}$ entre s y t opere.
- $(1 - r_{(s,u_i)}r_{(u_i,t)})$ es la probabilidad de que el camino de 2 aristas $\{s, u_i, t\}$ entre s y t falle.
- $r_{s,t}$ es la probabilidad que el enlace (s, t) opere y $1 - r_{s,t}$ la probabilidad de que falle.
- $(1 - r_{s,t}R')$ es la probabilidad de que todos los caminos de 2 aristas fallen simultáneamente.

La generalización de estas ideas para $D \geq 3$ y $|K| \geq 2$ es de complejidad NP -difícil [41, 44]. En particular si η es el grado máximo de la red, la identificación de todos los caminos relevantes aplicando al Algoritmo de Búsqueda en Profundidad requiere tiempo de cálculo de orden $O(\frac{k(k-1)}{2} \times \eta^D)$.

Los procedimientos de reducción ya referidos para el cálculo de $R_K(G)$ se generalizan para la medida de diámetro confiabilidad, lo que permite simplificar la topología de los grafos sobre los que se trabaja.

3.3.1. Procedimientos de reducción de grafos

Los procedimientos de reducción son válidos en métodos de cálculo exacto y aproximado, y permiten simplificar la topología de la red bajo estudio. Estas transformaciones de tiempo polinomial pueden afectar la distancia entre los nodos de la red y la cota de diámetro D asociada al problema. De esta forma los métodos presentados en el Capítulo 2 para $D = n - 1$ en general no son aplicables tal como fueron definidos, debiéndose considerar cambios en la cota D para preservar la medida $R_{K,D}(G)$.

Se menciona a continuación alguna de estas transformaciones aplicadas al caso s - t -terminal $R_{\{s,t\},D}(G)$ [46].

1. **Eliminación de enlaces irrelevantes:**

La *eliminación de enlaces irrelevantes* no afecta distancias entre nodos de forma que no modifica la cota de diámetro D : $R_{\{s,t\},D}(G) = R_{\{s,t\},D}(G - e)$, con e un enlace irrelevante de G . En el contexto de diámetro confiabilidad un enlace irrelevante es aquel que no pertenece a ningún D -path que conecte los nodos objetivo. En [47] se desarrolla un procedimiento para identificar aristas irrelevantes de forma eficiente.

2. Reducción path-terminal:

La *reducción path-terminal*, puede aplicarse cuando el grafo tiene un nodo terminal s de grado 1 conectado a un nodo u por medio de un camino P ,

$$P = \{(s, v_1), (v_1, v_2), \dots, (v_p, u)\},$$

$|P| = p + 1$, $\text{grado}(v_i) = 2$, $v_i \in P$ (ver Figura 3.5). Se cumple

$$R_{\{s,t\},D}(G) = R_{\{t,u\},D-p-1}(G') \prod_{k \in P} r_k,$$

donde $G' = (V - \{t, v_1, v_2, \dots, v_p\}, E - P)$.

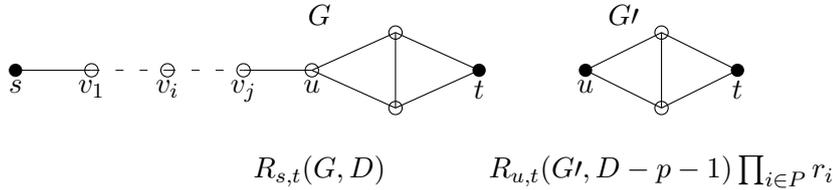


Figura 3.5: Reducción path-terminal

3. Reducción de enlaces en paralelo:

Sean e, e' dos enlaces en paralelo (que comparten ambos nodos extremos). La reducción de enlaces en paralelo reemplaza e y e' por un nuevo enlace f sin afectar D . Se cumple $r_f = r_e + r_{e'} - r_e r_{e'}$.

4. Caminos disjuntos:

La *transformación de caminos disjuntos* sólo se aplica si el conjunto E de G admite una partición en D -paths nodo disjuntos, P_1, P_2, \dots, P_m . La medida cumple

$$R_{\{s,t\},D}(G) = 1 - \prod_{k=1}^{k=m} (1 - \prod_{e \in P_k} r_e),$$

donde

- $\prod_{e \in P_k} r_e$ es la probabilidad de que el camino P_k esté operativo.
- $\prod_{k=1}^{k=m} (1 - \prod_{e \in P_k} r_e)$ es la probabilidad de que ninguno de los m caminos esté operativo.
- $R_{\{s,t\},D}(G)$ es la probabilidad de que al menos uno de los m caminos opere.

5. Aristas incidentes a un nodo terminal:

En el caso que todas las *aristas incidentes a un nodo terminal* están operativas puede aplicarse una reducción tal que si se contraen todos los nodos adyacentes siempre conexos al nodo terminal se obtiene un único nodo y se genera una subred G' más pequeña que verifica $R_{\{s,t\},D}(G) = R_{\{s,t\},D-1}(G')$.

3.3.2. Método de Factorización

El Método de Factorización aplica búsqueda recursiva sobre grafos definidos a partir de transformaciones aplicadas en el grafo en estudio. Para ello se basa en la dicotomía de estado de los enlaces y en procesos de reducción.

El método de factorización permite simplificar el recorrido por el espacio de vectores de estado de la red valiéndose de enlaces en estado no aleatorio. Un enlace e en estado de falla ($r_e = 0$) se elimina de la red mientras, que uno en estado operativo ($r_e = 1$) admite una contracción en la red.

La eliminación de un enlace tiene impacto negativo en la conectividad de la red dado que reduce el número de caminos factibles entre nodos objetivo. Este cambio puede repercutir de dos formas en el estado global de la red y determina cómo continúa el procedimiento recursivo. Sea $P \in P_{k_i, k_j}^D$ un camino factible de largo no mayor a D entre $k_i, k_j \in K$, $e \in P$. A partir de la eliminación de la arista e las alternativas son :

- Si todos los enlaces distintos de e que integran P fallan, entonces el par k_i, k_j no es conexo. Si no hay más caminos operativos entre k_i, k_j la red falla.
- Si al menos uno de los enlaces de P no falla, el proceso recursivo debe continuar

Por otro lado la contracción de una arista, aplicada a aristas que nunca fallan, también genera dos escenarios posibles:

- Si todos los enlaces distintos de e que integran P son perfectos, entonces el par k_i, k_j es conexo siempre.
- Si al menos uno de los enlaces de P no es perfecto, la búsqueda en el grafo debe continuar.

En base a estas ideas el teorema de factorización establece la siguiente relación recursiva:

$$R_{K,D}(G) = r_e R_{K,D}(G * e) + (1 - r_e) R_{K,D}(G - e) \quad (3.6)$$

donde e se denomina arista pivot, $G * e$ es la red que resulta de G luego de la contracción del enlace e y $G - e$ la red que resulta de G al eliminar e .

En el peor de los casos la implementación es de complejidad exponencial [44]. Sin embargo, la búsqueda puede consumir menos recursos como consecuencia de la eliminación o contracción de enlaces. Según se indica en [44], el método funciona correctamente para cualquier criterio de selección de la arista pivot, aunque el desempeño no es indiferente a dicho criterio.

El Algoritmo 11 resume la implementación del método de factorización. Los parámetros de entrada son el grafo G , el conjunto de nodos objetivo K , la confiabilidad elemental de los enlaces, \vec{r} , y la cota de diámetro D . El parámetro de salida es el valor exacto de la medida $R_{K,D}(G)$.

Asociado al Algoritmo 11 se consideran además dos algoritmos auxiliares: CONTRAER (Algoritmo 12) y ELIMINAR (Algoritmo 13) que aplican los mecanismos respectivos de contracción y reducción sobre el grafo. Ambos actualizan los parámetros recién detallados y definen

el subgrafo resultante de la transformación.

ALGORITMO 11 Factorización

ENTRADA:

grafo G ,
 nodos objetivo K ,
 confiabilidad de los enlaces \vec{r} ,
 cota de diámetro D .

SALIDA:

medida $R_{K,D}(G)$.

- 1: $R_{contraer} = 0$
 - 2: $R_{eliminar} = 0$
 - 3: Elegir arbitrariamente un enlace e , tal que $0 < r_e < 1$
 - 4: $[G * e, R_{contraer}] = CONTRAER(G, e)$
 - 5: **Si** $R_{contraer} = 1$ **entonces**
 - 6: $[G - e, R_{eliminar}] = ELIMINAR(G, e)$
 - 7: **Si no**
 - 8: $R_{contraer} = FACTO(G * e, D)$
 - 9: **Fin Si**
 - 10: FINAL
 - 11: $R_{K,D}(G) = r_e R_{contraer} + (1 - r_e) R_{eliminar}$
-

En las primeras dos líneas del Algoritmo 11 se definen las variables $R_{contraer}$ y $R_{eliminar}$ en las que se acumula el resultado de los Algoritmos CONTRAER y ELIMINAR respectivamente.

En la Línea 3 se elige la arista pivot e con confiabilidad elemental $r_e \in (0, 1)$. Un criterio posible de selección es $\{e : r_e = \max_{i \in E, r_i \in (0,1)} r_i\}$.

En la Línea 4 se aplica el Algoritmo CONTRAER. A partir de G y e retorna el grafo resultante de la contracción, $G * e$, y el parámetro $R_{contraer}$. En el Algoritmo principal, si $R_{contraer} = 1$ (Línea 5) la subred generada es un sistema determinista operativo y la búsqueda continúa en el Algoritmo ELIMINAR.

En la Línea 8, si la red $G * e$ admite más de un estado se aplica la llamada recursiva $FACTO(G * e, D)$.

Cuando los pasos recursivos culminan para ambos algoritmos auxiliares, como se indica en la Línea 11, se calcula la medida de confiabilidad de acuerdo con la Ecuación 3.6.

Para la implementación de los algoritmos auxiliares se definen los siguientes parámetros:

- $Conexo_{k_i, k_j}$: Variable de tipo Booleano. Toma Valor falso si no existe camino perfecto entre $k_i, k_j \in K$ de largo menor o igual a D y verdadero en otro caso.
- $Factible_p$: Variable de tipo Booleano. Dado un camino p , toma valor falso si p tiene al menos una arista en falla y verdadero si es un camino operativo.
- $N_{p_{k_i, k_j}}$: Valor entero. Indica el número de caminos entre $k_i, k_j \in K$ de largo menor o

igual a D .

- NE_p : Valor entero positivo. Indica el número de enlaces no perfectos en el camino p en G .
- NPC : Valor entero positivo. Indica número de pares conexos en el grafo G .
- $P(D)$: Unión de todos los caminos factibles entre k_i y k_j , P_{k_i, k_j}^D .
- $P(e)$: Conjunto de caminos en $P(D)$ que contienen a $e \in E$.

ALGORITMO 12 CONTRAER(G, e)

ENTRADA:

grafo G ,

arista pivot e .

SALIDA:

$R_{contraer}$.

- 1: **Para** $p = \{s, \dots, t\} \in P(e)$
 - 2: **Si** $Factible_p = Verdadero$ **entonces**
 - 3: Reducir el número de enlaces en falla: $NE_p = NE_p - 1$
 - 4: **Si** $Conexo_{s,t} = FALSO$ y $NE_p = 0$ **entonces**
 - 5: $Conexo_{s,t} = VERDADERO$
 - 6: $NPC = NPC + 1$
 - 7: **Si** $NPC = \frac{k(k-1)}{2}$ **entonces**
 - 8: $R_{contraer} = 1$
 - 9: **Fin Si**
 - 10: **Fin Si**
 - 11: **Fin Si**
 - 12: **Fin Para**
-

El Algoritmo 12 aplica el procedimiento de contracción del enlace e . A partir del grafo $G = (V, E)$ y la arista e actualiza los parámetros, retorna el grafo $G * e$ y el parámetro $R_{contraer}$.

Para todos los pathsets que contienen la arista e se repiten el bloque entre las Líneas 2 y 11. Dado que la contracción se aplica a una arista que siempre está operativa, para cada camino factible p que contiene la arista e se reduce el número de enlaces en estado de falla. En la Línea 4 se actualiza el número de enlaces en falla NE_p (Línea 3).

En la Línea 4, si, como consecuencia de la contracción, el camino p es siempre conexo ($NE_p = 0$) entonces p tienen todos sus enlaces perfectos. El número de pares conexos aumenta en una unidad: $NPC = NPC + 1$ (Línea 6).

En la Línea 7, si la variable NPC (número de pares conexos) coincide con $C_2^k = \frac{k(k-1)}{2}$ entonces todos los pares de nodos objetivo están conectados y la red opera. En este caso $R_{contraer} = 1$.

ALGORITMO 13 ELIMINAR**ENTRADA:**

grafo G ,
arista pivot e .

SALIDA:

$R_{eliminar}$.

- 1: **Para** $p = \{s, \dots, t\} \in P(e)$
- 2: **Si** $Factible_p = VERDADERO$ **entonces**
- 3: $Factible_p = FALSO$
- 4: $NPC = NPC - 1$
- 5: **Si** $NPC = 0$ **entonces**
- 6: $R_{eliminar} = 0$
- 7: **Fin Si**
- 8: **Fin Si**
- 9: **Fin Para**
- 10: $R_{eliminar} = FACTO(G - e, D)$

El Algoritmo 13 parte de la red $G = (V, E)$, el enlace e y retorna $R_{eliminar}$ y $G - e$. La aplicación itera tantas veces como caminos de G que contienen a e . Cada camino factible que contiene e se vuelve un camino no factible como consecuencia de la eliminación de e .

En la Línea 3 el estado del camino se cambia a no factible. En la Línea 4 se reduce el número de caminos factibles en una unidad, $NPC = NPC - 1$.

En la Línea 6, si no hay caminos que conecten ningún par del conjunto objetivo ($NPC = 0$), la red falla y $R_{eliminar} = 0$.

En la Línea 10, si la eliminación de e genera un grafo operativo el proceso continúa. Se aplica llamada recursiva sobre $G - e$.

El método de factorización es a lo sumo de complejidad computacional $O(2^m k^2 \delta^D)$ donde $|E| = m$, $|K| = k$ y δ es el grado máximo de la red [44], y su desempeño mejora si se consideran procedimientos de reducción para simplificar la topología del grafo en estudio.

3.3.3. Método Recursivo

En [48] se presenta una formulación recursiva para el cálculo exacto de la medida $R_{K,n-1}(G)$, que considera particiones de manera análoga al método aproximado *RVR* presentado en el Capítulo 2. A continuación se desarrolla el método para el caso genérico $D \leq n - 1$.

Dada una red $G = (V, E)$ con cota de diámetro D el método recorre el conjunto de vectores de estado de la red en función del estado de los componentes de un D -cutset de la red.

Sea C un D -cutset con probabilidad de falla Q_C . Se define una variable aleatoria Y_G , que vale 1 si el estado de la red es un estado de falla y que cumple:

$$Y_G^k = Q_C + (1 - Q_C)Y_{G_i}^k,$$

donde G_i es el subgrafo de G tal que las primeras $i - 1$ aristas de C fallan y la i -ésima opera, $Y_{G_i}^k$ es una realización de Y_{G_i} , siendo esta última la variable indicatriz asociada al estado de falla de G_i . El proceso se repite para $i = 1 \dots |C|$, y se cumple que Y_G^k es

$$Y_G^k = Q_C + (1 - Q_C) \sum_{i=1}^{|C|} 1_{\{V_C=i\}} Y_{G_i}^k$$

donde $1_{\{V_C=i\}}$ es independiente de Y_{G_i} , y V_C es una variable aleatoria discreta con función de masa de probabilidad

$$p(G_v) = \Pr(V = v) = \frac{1 - q_v}{1 - Q_C} \prod_{i=1}^v q_i. \quad (3.7)$$

Aplicando valor esperado en la definición de Y_G se define una fórmula recursiva sobre subgrafos de G :

$$E(Y_G) = Q_C + (1 - Q_C) \sum_{i=1}^{|C|} E(1_{\{V_C=i\}}) E(Y_{G_i}), \quad (3.8)$$

donde $E(1_{\{V_C=i\}}) = p(A_i)$, y $p(A_i) = \Pr(V_C = i) = \frac{r_i}{1 - Q_C} \prod_{j=1}^{i-1} q_j$, con $A_i \subseteq \Omega_G$. La fórmula recursiva para el cálculo exacto de $Q_{K,D}(G)$ es:

$$Q_{K,D}(G) = \begin{cases} 1, & \text{G no es } K\text{-conexo,} \\ 0, & |K| = 1, \\ Q_C + (1 - Q_C) \sum_{i=1}^{|C|} p(A_i) Q_{K,D}(G_i), & \text{en otro caso.} \end{cases} \quad (3.9)$$

El método parte de un grafo G , elige un D -cutset y define una partición del conjunto de configuraciones de G asociado a C , P_G . La recursión se aplica sobre cada componente de P_G , de forma que el árbol de la recursión tiene tantas ramificaciones como enlaces tiene C . Cada componente i -ésima de la partición es un conjunto de subgrafos de G condicionados a que la i -ésima arista del D -cutset opera y las $i - 1$ anteriores fallan. El proceso continúa hasta que se alcanzan estados deterministas del grafo.

ALGORITMO 14 ExactoRecurSivo**ENTRADA:**

grafo G ,
 nodos objetivo K ,
 confiabilidad de los enlaces \vec{r} ,
 cota de diámetro D .

SALIDA:

medida $Q_{K,D}(G)$.

```

1: Si  $|K| == 1$  entonces
2:    $S = 0$ 
3: Fin Si
4: Si  $G = (V, E)$  no es  $K$ -conexo entonces
5:    $S = 1$ 
6: Si no
7:    $C = \text{Corte}(G)$ 
8:    $Q = \prod_{i \in C, i=1}^{|C|} (q_i)$ 
9:    $pX = \text{ProbParticion}(C)$ 
10:   $S = Q$ 
11:  Para  $i = 1$  hasta  $|C|$ 
12:     $G_i; k \in C, k < i$  falla y  $i$  opera
13:     $aux = \text{ExactoRecurSivo}(GX, D)$ 
14:     $S = S + (1 - Q) * pX(i) * aux$ 
15:  Fin Para
16: Fin Si
17:  $Q_{K,D}(G) = S$ 

```

El Algoritmo 14 resume el proceso recursivo. Los parámetros de entrada son el grafo $G = (V, E)$, el conjunto de nodos objetivo, la confiabilidad elemental de los enlaces, \vec{r} y la cota de diámetro D . El parámetro que retorna es el valor exacto de la medida de anticonfiabilidad diámetro acotada $Q_{K,D}(G)$.

En las primeras 4 líneas Algoritmo 14 se evalúan casos base de la recursión.

En la Línea 1 se considera el caso en que la red está siempre operativa en función del cardinal del conjunto K . Si consta de un único nodo, $|K| = 1$, la red está siempre operativa y el algoritmo retorna $S = 0$.

En la Línea 4 se evalúa si la red está en estado de falla. Si éste es el caso el algoritmo retorna $S = 1$. Una alternativa para evaluar el estado de la red es aplicar el Algoritmo de Búsqueda en Anchura (BFS por sus siglas en inglés).

Si la red no es un sistema determinista se recorre el espacio de configuraciones posibles de la red, Ω_G . En la Línea 7 se define un D -cutset de G .

Dado C se define la partición P_G , del espacio de configuraciones de G y se calcula la probabilidad asociada a cada componente de la partición: $P_G = \{P_0, P_1, \dots, P_{|C|}\}$, con P_0 el conjunto de los subgrafos tal que C falla y P_i el conjunto de los subgrafos tal que opera la i -ésima arista de C y las $i - 1$ anteriores fallan.

En la Línea 8 se calcula la probabilidad de falla del D -cutset C :

$$Q = \prod_{i=1}^{|C|} q_i.$$

En la Línea 9 se calcula la probabilidad de cada componente de la partición de acuerdo con la Ecuación 3.7.

En la Línea 10 se inicializa el parámetro $S = Q$, en el que se acumula cada resultado de la recursión.

En el bloque de código indicado entre las Líneas 11 y 15 se repite para cada componente de la partición. Se define el grafo G_X a partir de G condicionado al elemento i de la partición. La llamada recursiva se aplica sobre G_X y el resultado, aux , incrementa S en $(1 - Q)p_X(i)aux$. El proceso se repite hasta que el subgrafo generado está en estado operativo o falla.

3.3.4. Enumeración de pathsets

Otra forma de interpretar la medida $R_{K,D}(G)$ se presenta en [44], donde los autores consideran la enumeración de todos los caminos de la red. Dada una red $G = (V, E)$ y el conjunto K de nodos objetivo, P_{k_i, k_j}^D es el camino de largo no mayor a D que conecta un par de nodos $k_i, k_j \in K$. Se define el evento $E(P_{i,j})$, “todos los componentes de P_{k_i, k_j}^D operan”, cuya probabilidad de ocurrencia, por la independencia de falla de los componentes, es:

$$\Pr(E(P_{i,j})) = \prod_{e \in P_{k_i, k_j}^D} r_e.$$

La medida de confiabilidad diámetro acotada es la probabilidad de que exista al menos un camino operativo entre todo par de nodos en K . En términos de eventos es la probabilidad de que ocurra al menos uno de los eventos $E(P_{i,j})$ para cada par $k_i, k_j, k_i \neq k_j$. Se cumple entonces:

$$R_{K,D}(G) = Prob \left(\bigcap_{k_i, k_j \in K} \left(\bigcup_{p \in P_{k_i, k_j}^D} E(P_{i,j}) \right) \right). \quad (3.10)$$

La no independencia entre eventos hace que el cálculo directo de la medida sea complejo. Como se observa en [44] la fórmula 3.10 evidencia que todo enlace que no pertenece P_{k_i, k_j}^D es irrelevante. En estos casos puede aplicarse reducción del grafo eliminando dichas aristas.

Otra forma sencilla de obtener de manera exacta la medida $R_{K,D}(G)$ es enumerar todas las configuraciones posibles de G y para cada una evaluar si corresponden a un estado operativo. La medida de confiabilidad diámetro acotada se obtiene como la suma de la probabilidad de que los pathsets de la red estén operativos. En [41] se detalla un algoritmo de búsqueda en profundidad (backtracking) para calcular la medida.

En el caso general los métodos exactos son costosos en tiempo y memoria, y de acuerdo con el interés que exista respecto a la exactitud de la medida puede considerarse su aplicación. En

otros contextos la complejidad del cálculo exacto no justifica su implementación y se aplican métodos aproximados que permitan estimar la medida con buena precisión a menor costo.

3.4. RVR: versión basada en muestreo multinomial

Los métodos de estimación presentados para el modelo clásico se adaptan al cálculo de la medida diámetro acotada. En esta sección se profundiza en el método *RVR* y se presenta una variante en el mecanismo de búsqueda que permite recorrer el espacio de vectores de estado de forma eficiente, para cualquier valor positivo del parámetro D [39]. Otros métodos pueden consultarse en [49] donde se comparan estimaciones de la medida de confiabilidad diámetro acotada.

En esta sección se presenta una nueva formulación del método *RVR*, propuesta por Cancela y El Khadiri [48]. El método aplica una búsqueda recursiva en el espacio de vectores de estado de G pero reduce el tiempo de ejecución respecto de la implementación original.

La aplicación original genera una muestra de tamaño N y define la variable aleatoria Y_G de acuerdo con la relación recursiva:

$$Y_G^k = Q_C + (1 - Q_C)Y_{G_i}^k,$$

donde Q_C es la probabilidad de falla del D -cutset C , G_i es un subgrafo de G tal que las primeras $i - 1$ aristas del D -cutset fallan y la i -ésima opera, y $Y_{G_i}^k$ es una realización de Y_G^k . El proceso se repite para $i = 1 \dots |C|$, de forma que Y_G^k es

$$Y_G^k = Q_C + (1 - Q_C) \sum_{i=1}^{|C|} 1_{\{V_C=i\}} Y_{G_i}^k. \quad (3.11)$$

con V_C una variable aleatoria discreta tal que $p(A_i) = \Pr(V_C = i) = \frac{r_i}{1-Q_C} \prod_{j=1}^{i-1} q_j$ y $A_i \subseteq \Omega_G$ es el i -ésimo componente de la partición P_G .

La nueva formulación recorre el espacio de vectores de estado Ω_G de G generando muestras en subconjuntos de Ω_G de acuerdo con un sorteo multinomial.

Al igual que en la formulación original se considera la función indicatriz $1_{\{V_C=i\}}$, con probabilidad $\Pr(1_{\{V_C=i\}} = 1) = p(A_i)$ y distribución Bernoulli($p(A_i)$).

En función de la probabilidad de falla de los enlaces y la partición asociada al conjunto de vectores de estado de la red se define la variable aleatoria N_i como el número estados a sortear en la i -ésima componente de una partición P_G de Ω_G . La variable N_i se distribuye Binomial de parámetros N y $p(A_i)$ tal que:

$$\sum_{i=1}^{|C|} N_i = N, \quad \sum_{i=1}^{|C|} p(A_i) = 1,$$

y el vector $\vec{N} = (N_1, N_2, \dots, N_{|C|})$ se distribuye Multinomial de parámetros N y $p = (p(A_1), p(A_2), \dots, p(A_{|C|}))$.

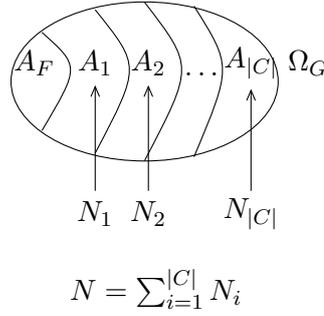


Figura 3.6: Muestreo multinomial

Definidas las variables, el estimador de $\hat{Q}_G = \frac{\sum_{i=1}^N Y_G^i}{N}$ puede expresarse como:

$$\hat{Q}_G = \frac{\sum_{i=1}^{|C|} \sum_{k=1}^{N_i} Y_{G_i}^k}{N}.$$

Sustituyendo Y_G^k por la expresión 3.11

$$\hat{Q}_G = \frac{NQ_C + (1 - Q_C) \sum_{i=1}^{|C|} \sum_{k=1}^{N_i} Y_{G_i}^k}{N}.$$

$$\hat{Q}_G = \frac{S1}{N} = \frac{NQ_C + (1 - Q_C) \sum_{i=1}^{|C|} S1(G_i, N_i)}{N}, \quad (3.12)$$

donde $S1(G_i, N_i) = \sum_{k=1}^{N_i} Y_{G_i}^k$.

Proposición 3.4.1 $\frac{S1}{N}$ es un estimador insesgado de $Q_{K,D}(G)$.

Demostración.

$$E(S1(G, N)) = E\left(NQ_C + (1 - Q_C) \sum_{i=1}^{|C|} S1(G_i, N_i)\right)$$

$$E(S1(G, N)) = NQ_C + (1 - Q_C) \sum_{i=1}^{|C|} E(S1(G_i, N_i)). \quad (3.13)$$

Dado que $S1(G_i, N_i)$ es una función de dos variables aleatorias, N_i e $Y_{G_i}^k$, puede aplicarse la propiedad $E(X) = E(E(X/Y))$ [50]. Se cumple:

$$E(S1(G_i, N_i)) = E\left(\sum_{k=1}^{N_i} Y_{G_i}^k\right) = E\left(E\left(\sum_{k=1}^{N_i} Y_{G_i}^k / N_i\right)\right) =$$

$$E\left(\sum_{k=1}^{N_i} E(Y_{G_i}^k)\right) = E\left(\sum_{k=1}^{N_i} E(Y_G^k/A_i)\right) = E(N_i E(Y_G/A_i))$$

$$E(Y_G/A_i)E(N_i) = E(Y_G/A_i)Np(A_i). \quad (3.14)$$

Por último, considerando las Ecuaciones 3.13, 3.14 y 3.8 se cumple:

$$E(S1(G, N)) = NQ_C + (1 - Q_C) \sum_{i=1}^{|C|} NE(Y/A_i)p(A_i)$$

$$E(S1(G, N)) = N(Q_C + (1 - Q_C) \sum_{i=1}^{|C|} E(Y/A_i)p(A_i))$$

$$E(S1(G, N)) = NE(Y_G)$$

$$E(S1(G, N)) = NQ_K(G),$$

y por lo tanto:

$$E\left(\frac{S1(G, N)}{N}\right) = Q_K(G)$$

QED

Se define además una variable auxiliar para estimar la varianza de la medida de confiabilidad, cuya distribución coincide con $\sum_{k=1}^N (Y_G^k)^2$. Dado el estimador insesgado de la varianza:

$$\frac{\sum_{k=1}^N (Y_G^k - \frac{S1}{N})^2}{N(N-1)}. \quad (3.15)$$

y la variable $S1$ 3.12 se define la variable $S2$:

$$S2(G, N) = NQ_C^2 + 2Q_C(1 - Q_C) \sum_{i=1}^{|C|} S1(G_i, N_i) + (1 - Q_C)^2 \sum_{i=1}^{|C|} S2(G_i, N_i) \quad (3.16)$$

Demostración.

$$\hat{V}_G = \frac{\sum_{k=1}^N (Y_G^k - \frac{S1}{N})^2}{N-1}$$

$$\frac{1}{N-1} \left(\sum_{k=1}^N \left(\frac{(Y_G^k)^2}{N} \right) - \left(\frac{S1}{N} \right)^2 \right). \quad (3.17)$$

Por otro lado:

$$(Y_G^k)^2 = (Q_C + (1 - Q_C)Y_{G_i}^k)^2.$$

Sustituyendo en la sumatoria resulta:

$$\sum_{k=1}^N (Y_G^k)^2 = \sum_{k=1}^N (Q_C + (1 - Q_C)Y_{G_i}^k)^2$$

$$NQ_C^2 + 2Q_C(1 - Q_C) \sum_{i=1}^{|C|} \sum_{k=1}^{N_i} Y^k(G_i) + (1 - Q_C)^2 \sum_{i=1}^{|C|} \sum_{k=1}^{N_i} (Y^k(G_i))^2 \quad (3.18)$$

Por último, sustituyendo la suma de las observaciones por $S1$ y la suma del cuadrado de las observaciones de acuerdo con 3.18:

$$S2(G) = NQ_C^2 + 2Q_C(1 - Q_C) \sum_{i=1}^{|C|} S1(G_i) + (1 - Q_C)^2 \sum_{i=1}^{|C|} S2(G_i)$$

QED

El estimador de la varianza 3.17 puede expresarse en función de ambas sumas, $S1$ y $S2$, como:

$$\hat{V}_G = \frac{1}{N(N-1)} \left[S2 - \left(\frac{S1}{N} \right)^2 \right] \quad (3.19)$$

Nótese que la forma de definir los parámetros $S1$ y $S2$ permiten acumular la información de la muestra durante el proceso recursivo evitando generar un vector con N entradas, lo que hace más eficiente el proceso de cálculo.

El Algoritmo 15 tiene como parámetros de entrada el grafo $G = (V, E)$, el conjunto de nodos objetivo K , la cota de diámetro D , la confiabilidad elemental de los enlaces, \vec{r} y el tamaño de muestra N . El algoritmo retorna la estimación de $Q_{K,D}(G)$, la probabilidad de falla de la red y su varianza.

Con S se denotan las sumatorias $S1$ y $S2$: $S = [S1, S2]$.

En la Línea 1 se consideran el caso trivial $|K| = 1$. Si este es el caso la red no requiere comunicación entre nodos especiales y la probabilidad de falla es nula: $S = [0, 0]$.

En la Línea 4 se evalúa si la red está siempre en estado de falla. Si esto se cumple la probabilidad de falla de la red es 1: $S = [N, N]$. Una alternativa para evaluar el estado de la red es un algoritmo de búsqueda en anchura.

Si el estado de la red no corresponde a ninguno de los casos recién mencionados se procede a la búsqueda de un D -cutset en función del que se determina una partición P_G de Ω_G .

En la Línea 7 se obtiene un D -cutset C .

En la Línea 8 se calcula la probabilidad de falla del D -cutset, como el producto de la probabilidad de falla de cada uno de sus componentes.

En la Línea 9 se calcula la probabilidad $p(A_i)$ asociada a la partición P_G de Ω_G en función del D -cutset. El i -ésimo elemento de π , se calcula de acuerdo con la siguiente expresión:

$$\pi_i = \frac{r_i \cdot \prod_{j=1}^{i-1} q_j}{1 - Q}.$$

En la Línea 10 se sortea una muestra multinomial de tamaño N con vector de probabilidad π [24]. Cada elemento N_i corresponde al tamaño de muestra que se obtendrá de la i -ésima

parte de P_G , tal que $\sum_{i=1}^{|C|} N_i = N$.

El bloque de Líneas 11-18 desarrolla la parte recursiva del algoritmo. En cada parte i de P_G donde el tamaño de muestra N_i es positivo se aplica la llamada recursiva del algoritmo sobre el subgrafo G_i de G . Como en la subred no se alteran distancias entre nodos sino solamente estados de los enlaces la cota D no cambia.

En la Línea 10 se inicializa el vector T en 0 en el que se van a acumular la suma de los resultados de cada llamada recursiva.

En la Línea 12 se obtiene el subgrafo G_i condicionado a la i -ésima partición.

ALGORITMO 15 RVRM

ENTRADA:

grafo G ,
 nodos objetivo K ,
 confiabilidad de los enlaces \vec{r} ,
 cota de diámetro D ,
 tamaño de la muestra N .

SALIDA:

estimación de la anticonfiabilidad: $\hat{Q}_{K,D}(G)$,
 estimación de la varianza: $\hat{V}_{K,D}(G)$.

- 1: **Si** $N == 0$ o $|K| == 1$ **entonces**
 - 2: $S = [0, 0]$
 - 3: **Fin Si**
 - 4: **Si** G no opera **entonces**
 - 5: $S = [N, N]$
 - 6: **Si no**
 - 7: $C = ConjCorte(G)$
 - 8: $Q = \prod_{i=1}^{|C|} (q_i)$
 - 9: $\pi = ProbParticion(C)$
 - 10: $T = [0, 0]$
 - 11: **Para** $i = 1$ **hasta** $|C|$
 - 12: $N_i = MuestraN(N, \pi_i)$
 - 13: $G_i; e_k \in C, k < i$ falla y e_i opera
 - 14: **Si** $N_i > 0$ **entonces**
 - 15: $T = T + RVRM(G_i, N_i, D)$
 - 16: **Fin Si**
 - 17: **Fin Para**
 - 18: $S = [NQ + (1 - Q)T(1), NQ^2 + 2Q(1 - Q)T(1) + T(2)(1 - Q)^2]$
 - 19: **Fin Si**
 - 20: $\hat{Q}_{K,D}(G) = \frac{S(1)}{N}$
 - 21: $\hat{V}_{K,D}(G) = \frac{1}{N(N-1)} \left[S(2) - \left(\frac{S(1)}{N} \right)^2 \right]$
-

A continuación se considera una aplicación del Algoritmo 15 sobre el grafo puente.

Ejemplo 3.4 Sea $G = (V, E)$ el grafo puente de la Figura 3.7 con nodos objetivo $K = \{s, t\}$. Los enlaces tal que $r_e = 1$ se representan con línea continua y los enlaces con $r_e \in (0, 1)$ se representan con línea punteada. El rectángulo en la parte superior izquierda de la Figura 3.7 indica el nombre de cada nodo del grafo y la forma de representar el estado de los enlaces. Cada enlace se denota por sus nodos extremos. Se aplica el Algoritmo 15 al grafo puente G con tamaño de muestra N y $D = 2$.

Se identifica un conjunto D -mincut para G , $C_1 = \{(s, u), (s, v)\}$. (Por simetría de G puede considerarse también el D -cutset $C_2 = \{(u, t), (v, t)\}$). Dado C_1 se define la partición de Ω_G , P_{G, C_1} la que se conforma de 3 elementos:

- P_0 contiene todos los subgrafos de G tal que $(s, u), (s, v)$ fallan (G_0 , Figura 3.7), $|P_0| = 2^3$.
- P_1 contiene todos los subgrafos de G tal que (s, u) opera (G_1 , Figura 3.7), $|P_1| = 2^4$.
- P_2 contiene todos los subgrafos de G tal que (s, u) falla y (s, v) opera (G_2 , Figura 3.7), $|P_2| = 2^3$.

En G_0 la búsqueda termina, $R_{K, D}(G_0) = 0$ y el algoritmo retorna $S = [0, 0]$. En los grafos G_1 y G_2 la recursión continúa, con tamaños de muestra N_1 y N_2 respectivamente. Se cumple $N = N_0 + N_1 + N_2$.

En G_1 se repite el proceso. En este caso el D -mincut es $C = \{(u, t), (v, t)\}$.

La partición para G_1 consta de 3 componentes:

- P_3 contiene todos los subgrafos de G_1 tal que $(u, t), (v, t)$ fallan (G_3 , Figura 3.7), $|P_3| = 2^2$.
- P_4 contiene todos los subgrafos de G_1 tal que (u, t) opera (G_4 , Figura 3.7), $|P_4| = 2^3$.
- P_5 contiene todos los subgrafos de G_1 tal que (u, t) falla y (v, t) opera (G_5 , Figura 3.7), $|P_5| = 2^2$.

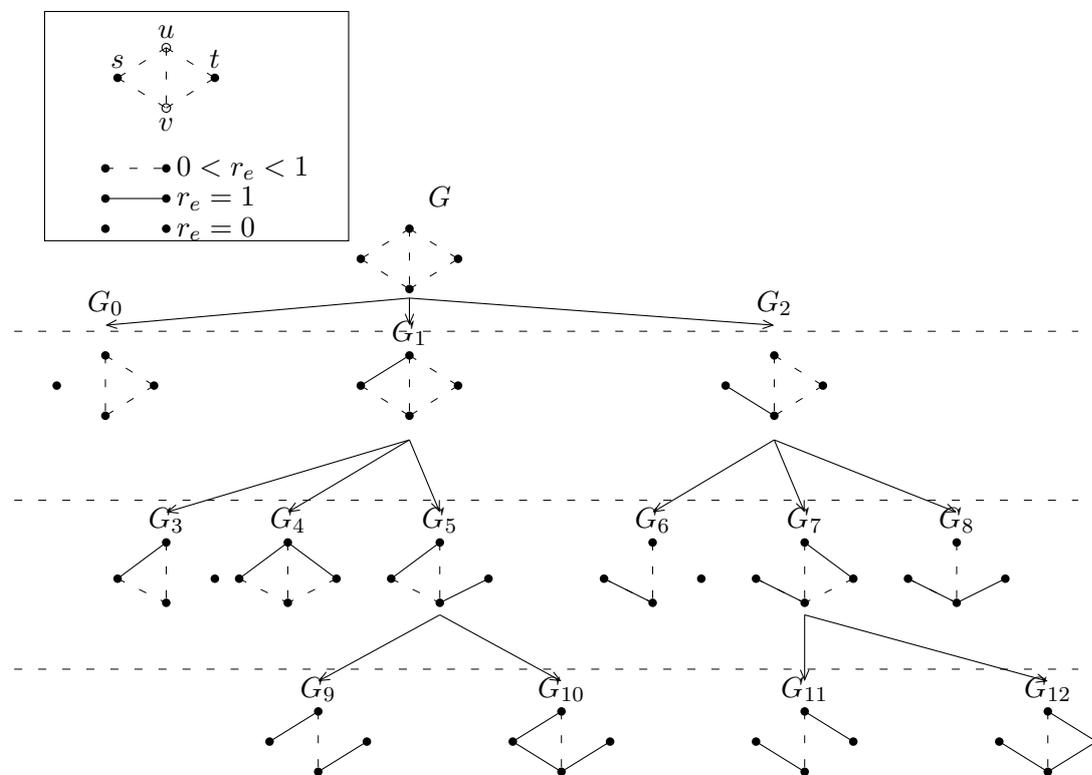


Figura 3.7: Ejemplo aplicación Algoritmo 15

En G_3 y G_4 la búsqueda termina, $R_{K,D}(G_3) = 0$, $R_{K,D}(G_4) = 1$, y el algoritmo retorna $S = [0, 0]$ y $S = [N_4, N_4]$ respectivamente. En el grafo G_5 la recursión continúa con tamaños de muestra N_5 . Se cumple $N_1 = N_3 + N_4 + N_5$. En esta instancia un D -cutset posible, integrado por $C = \{(s, v)\}$. La partición se conforma por:

- P_9 contiene todos los subgrafos de G_5 tal que (s, v) , falla (G_9 , Figura 3.7), $|P_9| = 2$.
- P_{10} contiene todos los subgrafos de G_5 tal que (s, v) opera (G_{10} , Figura 3.7), $|P_{10}| = 2$

En todos los casos la búsqueda termina, $R_{K,D}(G_9) = 0$, $R_{K,D}(G_{10}) = 1$, y el parámetro de salida vale $S = [0, 0]$, $S = [N_{10}, N_{10}]$. Se cumple $N_5 = N_9 + N_{10}$

En forma análoga se procede para G_2 . Una vez que se llega a las hojas del árbol de búsqueda (pasos base de la recursión) el procedimiento vuelve al punto de partida acumulando la información generada.

A modo indicativo, el valor real de la anticonfiabilidad elemental cuando $q_i = 0.9$, $\forall i \in E$ y $D = 2$ es 0.06688 y el valor estimado para una corrida específica del Algoritmo 15 con $N = 10^6$ es 0.06703.

3.5. Antecedentes

Con más de 50 años de investigación es vasta la literatura dedicada a la medida de confiabilidad. Sin embargo la generalización de este concepto surge en el año 2001 con la definición de la medida de confiabilidad diámetro acotada. En este sentido los aportes teóricos y empíricos para la medida de confiabilidad diámetro acotada son incipientes en relación al concepto predecesor.

En el año 2001 Cancela y Petingi [44] detallan el cálculo exacto para problemas donde $D = 1$, ó $D = 2$ y conjunto objetivo $K = \{s, t\}$. Obtienen además una cota superior e inferior para la medida de confiabilidad $R_{\{s,t\},D}$ computable en tiempo exponencial en D .

En el año 2004 se publica un documento dedicado al estudio de las propiedades de dominancia invariante, dada la importancia de dicho valor para los algoritmos de cálculo de la confiabilidad como a nivel teórico. En el documento se caracteriza la dominancia y se establece una regla de cálculo de la medida de confiabilidad diámetro acotada [51].

Otro resultado interesante del mismo año considera una subfamilia de grafos que se caracteriza por admitir una partición de los enlaces en cuatro subgrupos [52]. La evaluación de la confiabilidad para esta subfamilia requiere de $O(n^3)$ y $O(Dn^3)$ en espacio y tiempo respectivamente, cálculo eficiente para un problema que pertenece a la clase NP -difícil.

En el año 2005 se aplica la medida de confiabilidad V -terminal diámetro acotada para representar un modelo de búsqueda de información en redes P2P [42]. Se define una red G no dirigida integrada por N nodos entre los que existe comunicación. Como cada nodo no “conoce” a todos sus vecinos, la confiabilidad del enlace representa la probabilidad de intercambio. El largo de los caminos está dado por el número de nodos que un paquete de información visita durante la búsqueda. El interés en este contexto radica en maximizar la probabilidad de que todo nodo encuentre respuesta a las solicitudes de información que realiza, problema que equivale a maximizar la medida $R_{V,D}(G)$.

En el 2007 Leslie Murray presenta un trabajo donde se comparan cinco métodos Monte Carlo aplicados a la estimación de la confiabilidad diámetro acotada sobre tres redes elementales [49]. Allí se deduce que entre los métodos comparados, CMC , Dagger, Monte Carlo Permutación(Construcción-Destrucción), Entropía cruzada sobre CMC y permutación, el método Monte Carlo Permutación es el mejor método en la relación simplicidad-eficiencia.

Petingi presenta un resumen de propiedades combinatorias y computacionales de la medida $R_{K,D}(G)$ por extensión de conceptos ya desarrollados para la medida clásica $R_K(G) = R_{K,n-1}(G)$. Presenta también resultados y aplicaciones de la medida para redes que pueden modelarse como grafo dirigido o no dirigido [41, 53].

Los mismos autores en el mismo año proponen un método de evaluación recursiva exacta de la medida de confiabilidad diámetro acotada en base al principio de inclusión exclusión combinado con reducciones topológicas [46].

En 2010 Cancela, El Khadiri y Petingi proponen un algoritmo de tiempo polinomial que permite detectar y eliminar enlaces irrelevantes en la evaluación de la medida de confiabilidad

diámetro acotada s - t -terminal [47]. Esto permite reducir tiempos de cómputos tanto en métodos de cálculo exacto como aproximados.

3.6. Problemas abiertos

Son diversas las áreas en las que es posible profundizar como resultado de la definición reciente de la medida de confiabilidad diámetro acotada.

Una alternativa de trabajo es la profundización de métodos de cálculo exacto de la medida de confiabilidad en base a ideas de reducción de grafos y fórmulas recursivas [46, 52].

Otra línea de investigación es la identificación de sub-clases de grafos para las cuales existan métodos de evaluación en tiempo polinomial. En [52] se prueba que, en topologías completas cuyo conjunto de ejes puede particionarse en a lo sumo cuatro clases equiconfiabiles, $R_{\{s,t\},D}(G)$ puede calcularse en forma eficiente para D fijo.

Otra área interesante en contextos donde no es necesario conocer en forma exacta la medida es la obtención de cotas de la medida. En este caso resta profundizar en la identificación de métodos eficientes y generales [44].

En el método *RVR* puede abordarse el análisis de la profundidad de la recursión y la influencia del mecanismo de partición de Ω_G en la calidad de los resultados. El mecanismo de selección de D -cutset y la influencia que esta elección tiene en el desempeño del algoritmo es una dimensión ya considerada en [54, 55] para el caso $D = n - 1$.

Otra línea de trabajo tiene que ver con la obtención de cotas para el error relativo y la investigación de la precisión y robustez de los estimadores obtenidos mediante métodos recursivos de reducción de varianza. En particular la dependencia de los resultados en contexto de eventos raros y la identificación de cotas para el error relativo.

En los métodos aproximados considerar la partición de Ω_G como una combinación de D -cutset y D -pathset, evidenciar su impacto en el desempeño del algoritmo, ya presentado en [56] para el caso general.

Parte III

ANÁLISIS DE ROBUSTEZ

Capítulo 4

Métodos de cálculo: análisis de eficiencia y robustez

4.1. Introducción

En los Capítulos 2 y 3 se han presentado métodos de cálculo exacto y métodos aproximados para obtener la medida de confiabilidad clásica o diámetro acotada. Los métodos exactos permiten determinar el valor de la medidas con error 0 pero en determinados contextos la aplicación es prohibitiva dado el costo computacional asociado. Por otro lado los métodos aproximados, aunque también están condicionados por la dimensión del grafo, permiten estimar la medida de confiabilidad con un costo computacional inferior al que impone el cálculo exacto. Ahora bien, entre todos los métodos disponibles ¿cuál es el mejor y en función de qué parámetros se efectúa dicha comparación? ¿Qué parámetros resumen el desempeño de los métodos de cálculo?

En base al material presentado en [23], en este capítulo se presentan algunas herramientas que permiten evaluar los métodos en función del error en las estimaciones y del tiempo computacional que involucran. Usualmente se considera el método *CMC* como modelo de referencia respecto del cual se buscan mejoras.

El desempeño de dos o más métodos aplicados para resolver un mismo problema se evalúan los resultados empíricos puede compararse en términos de eficacia y eficiencia. Fijando el número de iteraciones una alternativa es comparar la varianza del estimador, que proporciona una medida de precisión de la estimación. Otra dimensión, asociada a la eficiencia, es el tiempo de cálculo necesario para obtener una estimación.

Si se comparan dos métodos, M_1 y M_2 en base a la varianza y el tiempo de cálculo puede considerarse la relación:

$$\frac{Var(Q_{M_1})T_{M_1}}{Var(Q_{M_2})T_{M_2}}, \quad (4.1)$$

donde $Var(Q_{M_i})$ y T_{M_i} son la varianza del estimador y el tiempo de cálculo del método M_i , $i = 1, 2$ respectivamente. Si M_2 tiene mejor desempeño que M_1 se obtiene un cociente mayor que 1 y esto implica que M_2 :

- obtiene la misma precisión que M_1 pero en menos tiempo, o bien,
- obtiene mejor precisión en el mismo tiempo, o bien,
- obtiene mejor resultado tanto en precisión como en tiempo de cómputo.

En la práctica, a efectos de comparar resultados, la disponibilidad de información y la posibilidad de homogeneizar criterios no siempre es posible. Actualmente no existe una biblioteca que concentre la información y proporcione casos de prueba de forma de comparar con éstos sistemáticamente, y, aunque existiese un vasto registro, el uso de diferentes computadores afectan los tiempos de cálculo y la comparación pierde validez.

A continuación se detallan indicadores que permiten resumir y comparar el desempeño de métodos aproximados, proporcionando una medida de calidad de desempeño de procedimientos de interés. En la Sección 4.2 se presenta la función de cobertura que permite evaluar la calidad de las estimaciones en función de los intervalos de confianza del estimador. En la Sección 4.3 se presentan propiedades asintóticas de los estimadores y características que dan cuenta de su robustez en función de la probabilidad de falla de los enlaces de la red.

En general, en teoría de eventos raros, el parámetro de rareza se denota con ϵ , donde $\epsilon \rightarrow 0$. En esta sección el parámetro de rareza se denota con q

4.2. Cobertura e intervalos de confianza

La función de cobertura, presentada por Schruben [57] en 1980, es una herramienta que permite evaluar la robustez de un estimador. Dado el valor real del parámetro que se desea estimar, la estimación correspondiente y su varianza deben generar intervalos de confianza cuya cobertura coincida con la cobertura teórica. Una forma de evaluar esta propiedad es generar intervalos de confianza independientes y comparar la frecuencia con que el parámetro real pertenece al intervalo teórico. Schruben propone la función de cobertura que permita evaluar la robustez de los intervalos de confianza estimados para cada probabilidad de cobertura $(1 - \alpha) \in [0, 1]$ [57].

Definición 4.2.1 Dado un estimador por intervalo $[L(X), U(X)]$ de un parámetro θ , **el coeficiente de confianza** de $[L(X), U(X)]$ es el ínfimo de las probabilidades de cobertura: $\inf_{\theta} \Pr(\theta \in [L(X), U(X)])$.

Observación: como no se conoce el verdadero valor del parámetro θ sólo puede garantizarse probabilidad de cobertura igual al ínfimo.

Sean $Q_{K,D}(G)$, \hat{Q} , \hat{V}_Q el valor real del parámetro, la estimación y la varianza del estimador respectivamente. Bajo el supuesto de que \hat{Q} se distribuye normal, el intervalo de confianza de nivel α para $Q_{K,D}(G)$ es de la forma:

$$IC_{\alpha} = \left[\hat{Q}_G - z_{\frac{\alpha}{2}} \sqrt{\hat{V}_{Q_G}}, \hat{Q}_G + z_{\frac{\alpha}{2}} \sqrt{\hat{V}_{Q_G}} \right],$$

donde $z_{\frac{\alpha}{2}} = \Phi^{-1}(1 - \frac{\alpha}{2})$ y Φ la función de distribución Normal(0, 1). Se cumple:

$$\Pr(Q_{K,D}(G) \in IC_{\alpha}) = 1 - \alpha = \nu$$

Se define una variable aleatoria ν^* :

$$\nu^* = \inf\{\nu \in [0, 1] : Q_G \in IC_{\alpha}\},$$

que se interpreta como el intervalo de confianza más pequeño tal que el parámetro real está en el intervalo IC_{α} .

Se cumple $F_{\nu^*}(\nu) = \Pr(\nu^* \leq \nu) = \nu$, con $\nu \in [0, 1]$.

Si se despeja ν del intervalo de confianza resulta que:

$$Q_G \in \left[\hat{Q}_G - z_{\{\frac{1-\nu}{2}\}} \sqrt{V_{\hat{Q}_G}}, \hat{Q}_G + z_{\{\frac{1-\nu}{2}\}} \sqrt{V_{\hat{Q}_G}} \right].$$

$$|\hat{Q}_G - Q| \leq z_{\{\frac{1-\nu}{2}\}} \sqrt{V_{\hat{Q}_G}}$$

$$\frac{|\hat{Q}_G - Q|}{\sqrt{V_{\hat{Q}_G}}} \leq z_{\{\frac{1-\nu}{2}\}}$$

$$\Phi \left(\frac{|\hat{Q}_G - Q|}{\sqrt{V_{\hat{Q}_G}}} \right) \leq \frac{\nu + 1}{2}$$

$$2\Phi \left(\frac{|\hat{Q}_G - Q|}{\sqrt{V_{\hat{Q}_G}}} \right) - 1 \leq \nu$$

Luego, como $2\Phi \left(\frac{|\hat{Q}_G - Q|}{\sqrt{V_{\hat{Q}_G}}} \right) - 1 \leq \nu$, se cumple que:

$$2\Phi \left(\frac{|\hat{Q}_G - Q|}{\sqrt{V_{\hat{Q}_G}}} \right) - 1 = \inf\{\nu \in [0, 1] : Q_G \in IC_{\alpha}\},$$

y por lo tanto:

$$\nu^* = 2\Phi \left(\frac{|\hat{Q}_G - Q|}{\sqrt{V_{\hat{Q}_G}}} \right) - 1.$$

Proposición 4.2.2 *La variable aleatoria ν^* tiene distribución Uniforme(0,1).*

Demostración.

$$\begin{aligned}
F_{\nu^*}(\nu) &= \Pr(\nu^* \leq \nu) = \text{Prob} \left(2\Phi \left(\frac{|\hat{Q}_G - Q|}{\sqrt{\hat{V}_{\hat{Q}_G}}} \right) - 1 \leq \nu \right) = \\
&= \text{Prob} \left(\Phi \left(\frac{|\hat{Q}_G - Q|}{\sqrt{\hat{V}_{\hat{Q}_G}}} \right) \leq \frac{\nu + 1}{2} \right) = \\
&= \text{Prob} \left(\frac{|\hat{Q}_G - Q|}{\sqrt{\hat{V}_{\hat{Q}_G}}} \leq \Phi^{-1} \left(\frac{\nu + 1}{2} \right) \right) = \\
&= \text{Prob} \left(\frac{|\hat{Q}_G - Q|}{\sqrt{\hat{V}_{\hat{Q}_G}}} \leq \Phi^{-1} \left(1 - \frac{1 - \nu}{2} \right) \right) = \\
&= \text{Prob} \left(\frac{|\hat{Q}_G - Q|}{\sqrt{\hat{V}_{\hat{Q}_G}}} \leq z_{\{\frac{1-\nu}{2}\}} \right) = \\
&= \text{Prob} \left(z_{\{\frac{1-\nu}{2}\}} \leq \frac{\hat{Q}_G - Q}{\sqrt{\hat{V}_{\hat{Q}_G}}} \leq z_{\{\frac{1-\nu}{2}\}} \right) = \\
&= 2 \left(\Phi(z_{\{\frac{1-\nu}{2}\}}) - \frac{1}{2} \right) = 2 \left(1 - \frac{1 - \nu}{2} - \frac{1}{2} \right) = \nu
\end{aligned}$$

QED

Dado un intervalo de confianza con nivel de confianza $\alpha = 1 - \nu$, si el método de estimación es correcto la cobertura obtenida coincide con la teórica ($F_{\nu^*}(\nu) = \nu$). Cuando se registran diferencias existen dos posibilidades, $F_{\alpha^*}(\alpha) < \alpha$, la cobertura es menor a la teórica y puede arribarse a conclusiones erróneas, o, $F_{\alpha^*}(\alpha) > \alpha$. En este caso el procedimiento es ineficiente y la cobertura deseada podría lograrse con un número de iteraciones menor.

La función de cobertura se define para todos los coeficientes de confianza entre 0 y 1, aunque bajos niveles de confianza no tienen interés desde el punto de vista práctico son importantes por su aporte a la función de cobertura.

En el contexto de confiabilidad la robustez del parámetro se mide en función de la rareza de los eventos de falla. Si q es la probabilidad de falla entonces el método de estimación es robusto si el intervalo estimado tiene cobertura mayor o igual a la cobertura deseada cuando $q \rightarrow 0$.

4.2.1. Función de cobertura: análisis empírico

La distribución empírica de la función de cobertura es de la forma:

$$G_{\nu^*}(\nu) = \frac{1}{m} \sum_{j=1}^m 1_{\{\nu_j^* \leq \nu\}},$$

donde $\nu_j^*, j = 1, 2, \dots, m$ son los valores observados de los coeficientes de confianza (ν) aleatorios de una muestra de tamaño m y la función $1_{\{\nu_j^* \leq \nu\}}$.

Dada la distribución empírica el desempeño del estimador se evalúa comparando la función simulada con la distribución Uniforme(0,1), aplicando la prueba estadística de bondad de ajuste de Kolmogorov-Smirnov [58].

4.3. Propiedades asintóticas

En la estadística clásica es usual evaluar estimadores de acuerdo con el número de réplicas, considerando dos propiedades asintóticas, consistencia y eficiencia. Ambas permiten evaluar la calidad del estimador en función del tamaño de muestra de forma que cuando éste crece la información proporcionada mejora y el estimador debería estar arbitrariamente cerca del valor real con alta probabilidad.

En el contexto de confiabilidad y fallas catalogadas como evento raro es de interés que la calidad de las estimaciones no se vea afectada cuando la probabilidad de falla se vuelve más pequeña. Formalmente, es deseable que el error relativo de la estimación de la medida de confiabilidad de G se mantenga acotado cuando $q \rightarrow 0$, donde q es el parámetro de rareza y coincide con la probabilidad de falla de los enlaces de G .

Junto con la precisión del estimador y la corrección de las estimaciones en contextos de eventos raros, debe tenerse en cuenta el esfuerzo computacional para generar el valor de la medida. La eficiencia del método depende del costo computacional requerido para generar la estimación, y, al incorporar esta dimensión, la calidad global del estimador es un compromiso entre precisión y esfuerzo computacional.

Los indicadores que se presentan a continuación son: error relativo (ER) que se centra puntualmente en la precisión numérica del resultado, eficiencia relativa ($EfRel$) que incluye el tiempo de cómputo necesario, y aproximación normal de intervalos de confianza.

- **Error relativo y Error relativo acotado**

Si \hat{R}_n es el estimador de R , calculado a partir de n simulaciones, una propiedad deseable es que el error relativo no se vea afectado por la baja probabilidad de las fallas en el sistema. Formalmente, es lo que se conoce como error relativo acotado, BRE por sus siglas en inglés, donde el error relativo se define como la mitad de la amplitud del intervalo de confianza de R_n :

$$ER = z_\alpha \frac{\sigma_n}{\hat{R}_n} \quad (4.2)$$

con σ_n^2 la varianza del estimador \hat{R}_n y $z_\alpha = \Phi^{-1}(1 - \alpha)$ con Φ la distribución normal estándar.

Se dice que un estimador es BRE si ER tiene cota superior cuando $q \rightarrow 0$. La propiedad BRE es difícil de verificar en la práctica por lo que se considera una optimalidad asintó-

tica en donde se considera eficiencia logarítmica respecto al parámetro de rareza [23]:

$$\lim_{q \rightarrow 0} \frac{\ln E[\hat{R}_n^2]}{R} = 2.$$

La cantidad bajo el límite es válida porque $Var(\hat{R}_n) \geq 0$, entonces $E(\hat{R}_n^2) \geq R^2$ y $\ln E(\hat{R}_n^2) \geq 2 \ln R$.

■ **Eficiencia y eficiencia relativa**

Cuando se analiza un estimador se genera un compromiso en esfuerzo computacional y precisión con el fin de obtener estimaciones de calidad a baja costo.

Se define la eficiencia de un estimador \hat{R}_n como una función inversa del tiempo computacional requerido para el cálculo: si n es el tamaño muestral, σ_n^2 la varianza del estimador y t_n el tiempo computacional promedio de cómputo entonces la eficiencia es

$$Ef = \frac{1}{\sigma_n^2 t_n}. \quad (4.3)$$

Al igual que en el caso del error se define un indicador relativo de eficiencia dada por la expresión $EfRel = \frac{R^2}{\sigma_n^2 t_n}$. Si además se cumple $EfRel > d$ cuando $q \rightarrow 0$ entonces la eficiencia relativa está acotada respecto al parámetro de rareza. Esto significa que el tiempo y la varianza no se desvirtúa y el estimador es robusto frente a eventos raros.

■ **Intervalos de confianza y aproximación normal acotada**

Otro aspecto a tener en cuenta además de los estimadores puntuales de la confiabilidad es la amplitud de los intervalos de confianza respectivos y su comportamiento en relación al parámetro de rareza q . Si un intervalo de confianza es normal acotado, *BNA* por sus siglas en inglés, entonces se necesitan un número fijo de iteraciones para que el intervalo de confianza tenga el mismo error sin importar el nivel de rareza de q .

Definición 4.3.1 Un estimador \hat{R}_n tiene **Aproximación Normal Acotada** si el cociente $\frac{\rho}{\sigma^3}$ permanece acotado cuando $q \rightarrow 0$ donde $\rho = E(X^3)$, con $X \sim Normal(0,1)$.

Se define también una versión generalizada de la aproximación:

Definición 4.3.2 Un estimador \hat{R}_n tiene **Aproximación Normal Acotada** si existe δ tal que $E[|X - R|^{2+\delta}] \sigma - (2 + \delta)$ permanece acotado cuando $q \rightarrow 0$.

El objetivo es generar intervalos de confianza cuya amplitud no se desvirtúa cuando aumenta la rareza del evento. Esta propiedad es condición suficiente pero no necesaria y asegura que se necesitan un número fijo de iteraciones para generar un intervalo de confianza con error fijo e independiente del nivel de rareza del evento.

Capítulo 5

Diagnóstico de robustez de *RVR*

En este capítulo se presenta el estudio de robustez del método *RVR* aplicado a la estimación de la medida $Q_{K,D}(G)$ en redes altamente confiables. El análisis se basa en la propiedad *BRE* (error relativo acotado) y en la función de cobertura.

La elección del método radica en su buen desempeño computacional para estimar la medida de confiabilidad en todo tipo de redes, con resultados destacados en redes altamente confiables [39], [48].

En la Sección 5.1 se presentan los resultados numéricos obtenidos para redes usualmente consideradas en la literatura y para la red de transporte de la empresa telefónica estatal ANTEL. Puntualmente se obtienen la estimación de la medida de anticonfiabilidad $Q_{K,D}(G)$, su varianza, el error relativo e indicadores que resumen la reducción de tiempo y varianza del método *RVR* respecto a *CMC*.

A partir de las estimaciones e indicadores obtenidos se analiza el método *RVR* en términos de precisión y eficiencia para distintos valores de anticonfiabilidad elemental q y cota de diámetro D .

En la Sección 5.2 se presenta un diagnóstico empírico de robustez basado en la función de cobertura. En esta instancia se consideran tres grafos de prueba y se simula la función de cobertura. El objetivo es determinar si los intervalos de confianza estimados son correctos independientemente de la magnitud de q .

Los resultados indican que la magnitud que el error relativo obtenido por el método *RVR* se ve afectado por la magnitud de la probabilidad de falla de los enlaces.

En la Sección 5.3 se detalla el estudio asintótico del error relativo para un grafo sencillo. En función de los resultados obtenidos en las Secciones 5.1 y 5.2, el estudio empírico se completa con el cálculo del límite del error relativo cuando $q \rightarrow 0$.

Se concluye que el método *RVR*, en todos los casos de prueba elegidos, es más eficiente que *CMC*. Respecto de la evolución del error relativo en función de q , *RVR* genera resultados diferenciados según la topología sobre la que se aplica. En algunos casos *ER* es decreciente cuando disminuye q y en otros creciente.

A partir de los resultados y el estudio asintótico, se concluye que la versión implementada en

esta tesis, en el caso general, no es robusta en contexto de eventos raros.

5.1. Resultados numéricos: estudio de casos

Esta sección presenta una aproximación a la eficiencia y robustez del método RVR considerando casos de prueba.

En la Subsección 5.1.1 se detalla la implementación del método RVR basado en muestreo multinomial y se explica brevemente la metodología de trabajo en el cluster FING.

En la Subsección 5.1.2 se detallan los casos de prueba elegidos.

En la Subsección 5.1.3 se define la notación utilizada para presentar los resultados numéricos. Por último en la Subsección 5.1.4, para cada caso de estudio, se presentan y analizan los valores obtenidos.

5.1.1. Implementación

Implementación de RVR

Todas las estimaciones se obtienen mediante la implementación del método RVR basado en muestreo multinomial (Ver Algoritmo 15). La partición del espacio de vectores de estado de G se define en función de un conjunto de corte de la red.

El estado global de la red se evalúa mediante un algoritmo de búsqueda en profundidad.

Por otra parte, no se aplicaron procedimientos de reducción de grafos. En particular se descartaron procedimientos de reducción serie-paralelo por la alteración que generan en la distancia entre nodos. La reducción afecta la topología de forma que la medida de confiabilidad diámetro acotada del grafo resultante no es equivalente a la medida del grafo original.

Ejemplo 5.1 Sea $G = (V, E)$ el grafo de la Figura 5.1 con $K = \{s, t\}$, $D = 1$. La confiabilidad de los enlaces es tal que: el enlace (s, u) nunca falla ($r_{(s,u)} = 1$) y los enlaces en $V - (s, u)$ operan con probabilidad r , $0 \leq r \leq 1$.

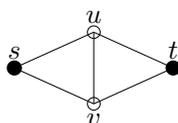


Figura 5.1: Ejemplo: reducción afecta distancias

Si se aplican procedimientos de reducción, entonces la arista (s, u) admite una contracción, de forma que s y u se fusionan en un único nodo. Como consecuencia de este cambio, los nodos s y v quedan conectados por dos aristas en paralelo que pueden reducirse a una única arista, con confiabilidad elemental $2r - r^2$. El grafo resultante es el de la Figura 5.2.

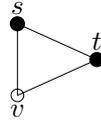


Figura 5.2: Ejemplo: reducción afecta distancias

Mientras en el grafo de la Figura 5.1 $Q_{\{s,t\},1}(G) = 1$ porque no existe camino de largo 1 que conecte los nodos objetivo, en el grafo de la Figura 5.2 la red está operativa cuando (s, t) opera, y $Q_{\{s,t\},1}(G) = 1 - r$. La reducción no preserva la medida de anticonfiabilidad y su aplicación no es válida.

Selección de D -cutset

Un D -cutset puede obtenerse en base al Algoritmo de Ford-Fulkerson, pero no todos los D -Cutset provienen de este algoritmo. La implementación de Ford-Fulkerson permite identificar un D -cutset de la red. En el caso s - t -terminal la aplicación del algoritmo es directa. En el caso K -terminal se eligen dos nodos del conjunto objetivo y se busca un D -cutset que los separe. Si no existe corte para el par elegido se continúa con otro par de nodos hasta encontrar un corte que separe al menos dos nodos de K .

En ambos casos, s - t -terminal y K -terminal, si el algoritmo Ford-Fulkerson no encuentra ningún conjunto de corte separador se evalúa si existe un corte alternativo. Se identifica el conjunto conformado por todas las aristas del grafo con probabilidad de falla positiva y se evalúa si la eliminación de ese conjunto impide la comunicación entre nodos objetivo. Esta forma de selección del conjunto de corte permite identificar, si existe, un D -cutset que no separa la red, pero que impide la comunicación por caminos de largo menor o igual a D .

Ejemplo 5.2 Sea $G = (V, E)$ el grafo de la Figura 5.3, con nodos objetivo $K = \{s, t\}$ y $D = 2$. La confiabilidad elemental es la siguiente: $r_e = 1$ para $e \in \{(s, u), (u, v), (v, t)\}$ y $r_e = r, 0 < r \leq 1$, para $e \in \{(s, v), (u, t)\}$.

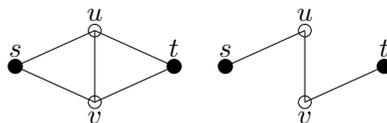


Figura 5.3: Ejemplo: conjunto de corte

La aplicación de Ford-Fulkerson no identifica un D -cutset. Aunque se eliminen los enlaces con confiabilidad r , $\{(s, v), (u, t)\}$ (Ver Figura 5.3), los enlaces $\{(s, u), (u, v), (v, t)\}$ operan siempre y según Ford-Fulkerson no existe conjunto de corte que separe s y t . Sin embargo como $D = 2$, la eliminación de $\{(s, v), (u, t)\}$ interrumpe la comunicación entre

s y t por caminos de largo 2. El conjunto $\{(s, v), (u, t)\}$ es un D -cutset y la anticonfiabilidad de la red es $Q_{\{s,t\},2} = (1 - r)^2$.

Infraestructura de cómputo

Los algoritmos se implementan en MATLAB, versión 7.11.0.584 y la infraestructura utilizada es el cluster de Facultad de Ingeniería.

El cluster FING es una infraestructura de cómputo de alto desempeño que permite resolver problemas complejos que requieren poder de cómputo¹. Entre otras cosas posibilita:

- procesar múltiples trabajos en forma simultánea.
- aplicar procesos recursivos que requiere capacidad de almacenamiento.
- aplicar procesos de cálculo exacto con tiempos de cómputo elevado.

Dos aplicaciones complementarias que se utilizaron para ejecutar los trabajos en el cluster son:

- PuTTY²: PuTTY es un cliente SSH (Secure SHell, en español: intérprete de órdenes segura) y Telnet (TELEcommunication NETwork: protocolo de red que permite acceder a otra en forma remota a través de una red) con licencia libre, que permite acceder a un servicio remoto en otro computador.
- PSFTP: protocolo SSH de red para transferencia y manipulación segura de datos, conocido como SFTP o Secure File Transfer Protocol. Permite el intercambio de archivos en forma segura mediante conexión remota entre computadores.

En el marco de esta tesis, cada trabajo que se procesa en el cluster se envía utilizando PuTTY de acuerdo con un script de extensión sh³. (En el Anexo 8 se presenta un script como ejemplo). Cada trabajo, una vez enviado, queda en lista de espera y se ejecuta en función de los recursos disponibles y los recursos requeridos por el algoritmo a implementar.

En cada tarea se especifica el tiempo máximo de cómputo (walltime), el archivo de Matlab a ejecutar y la dirección de correo electrónico de notificación. Si un trabajo alcanza el tiempo máximo especificado el cluster interrumpe la ejecución. En todos los casos se obtiene una notificación por correo electrónico que confirma que un trabajo fue enviado o que el proceso de cómputo finalizó.

La aplicación PSFTP permite intercambiar archivos entre la computadora del usuario y el cluster de manera segura.

5.1.2. Casos de estudio

A continuación se detallan las topologías consideradas para el estudio experimental, que coinciden con los casos frecuentemente citados en la bibliografía consultada. Se considera además la red de ANTEL como un caso de aplicación a una red real actual.

¹<http://www.fing.edu.uy/cluster>.

²<http://www.putty.org>

³extensión de archivos ejecutables en linux

1. Caso s - t -terminal en Arpanet. Valores de $q \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}\}$.
2. Caso s - t -terminal en Dodecahedro. Valores de $q \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}\}$.
3. Caso s - t -terminal en Dodecahedro, con q variando en dos valores, q_h y q_r .
4. Caso K -terminal en Grilla $n \times n$, $n \in \{3, 5, 6\}$. Valores de $q \in \{10^{-3}, 10^{-6}\}$.
5. Caso s - t -terminal en grafo completo K_n , $n = \{4, 5, 6, 7, 8\}$. Valores de $q \in \{10^{-1}, 10^{-2}\}$.
6. Caso red real: Red de transporte SDH (fibra óptica) de ANTEL. Se evalúan dos subgrafos, regiones este y oeste, donde:
 - $K = \{v \in V; \text{grado}(v) \geq 3\}$,
 - $q \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}\}$.

El Cuadro 5.1 resume las dimensiones que caracterizan a los distintos grafos.

Los valores de D considerados son tales que $D_{op} \leq D \leq n - 1$. El valor D_{op} es el mínimo valor de cota diámetro tal que para cualquier $D < D_{op}$ la red siempre falla:

$$Q_{K,D}(G) = 1, \forall D < D_{op}.$$

Por otro lado toda cota de diámetro D tal que $D \geq n - 1$ genera la medida clásica de arista confiabilidad de la red:

$$Q_{K,D}(G) = Q_{K,n-1}(G) = Q_K(G), \forall D \geq n - 1,$$

pues los nodos de la red están separados a lo sumo por $n - 1$ enlaces.

Cuadro 5.1: Dimensiones de los grafos de prueba.

Grafo	$ V $	$ E $	$ K $	D_{op}
K_n	n	$\frac{n(n-1)}{2}$	2	1
Grilla $n \times n$	n^2	$2n(n-1)$	4	$2(n-1)$
Arpanet	21	26	2	5
Dodecahedro	20	30	2	4
ANTEL: Red Este	58	65	13	13
ANTEL: Red Oeste	53	57	8	18

5.1.3. Notación

- \hat{Q}_{RVR} : estimación de la medida de anticonfiabilidad $Q_{K,D}(G)$ aplicando el método RVR , tal que $E(\hat{Q}_{RVR}) = Q_{K,D}(G)$.
- \hat{V}_{RVR} : estimación de la varianza de \hat{Q}_{RVR} .

- V_{CMC} : valor de la varianza para el método CMC : $V_{CMC} = \frac{\hat{Q}_{CMC}(1-\hat{Q}_{CMC})}{N-1}$.
- $ER = \frac{\sqrt{\hat{V}_{RVR}}}{\hat{Q}_{RVR}}$, es el error relativo del método RVR .
- $V_{CMC/RVR} = \frac{\hat{V}_{CMC}}{\hat{V}_{RVR}}$. Si $V_{CMC/RVR} = k$ entonces el método RVR tienen varianza k veces menor que CMC .
- $W_{CMC/RVR} = \frac{t_{CMC}\hat{V}_{CMC}}{t_{RVR}\hat{V}_{RVR}}$. Cociente de la varianzas por tiempos de cómputos de ambos métodos, donde t_{RVR} es tiempo de cómputo necesario para generar la estimación \hat{Q}_{RVR} y t_{CMC} es el tiempo de cómputo de la estimación \hat{Q}_{CMC} . Es un indicador global de cuánto mejora la eficiencia del método RVR respecto de CMC para la misma cantidad de iteraciones N .

Respecto del método CMC es importante precisar cómo se obtienen la varianza y el tiempo de ejecución correspondiente. Dado que no tiene buen desempeño en redes altamente confiables, para obtener una estimación de la medida de anticonfiabilidad con buena precisión sería necesario generar un número impracticable de iteraciones.

Sin embargo, puede efectuarse la comparación con RVR para un mismo tamaño de muestra. El tiempo de cómputo de CMC , t_{CMC} , es lineal en el número de iteraciones N y por lo tanto el cociente $K_{CMC} = \frac{N}{t_{CMC}} \simeq \frac{N-1}{t_{CMC}}$ es constante (para una misma máquina).

Para comparar el desempeño de CMC con RVR se considera la siguiente relación:

$$V_{CMC}t_{CMC} = \frac{Q_{K,D}(G)(1 - Q_{K,D}(G))}{N - 1}t_{CMC} = \frac{Q_{K,D}(G)(1 - Q_{K,D}(G))}{K_{CMC}}.$$

donde el producto $V_{CMC}t_{CMC}$ no depende de N ni de t_{CMC} y el valor K_{CMC} puede determinarse experimentalmente para un número pequeño de iteraciones. El valor de la medida de anticonfiabilidad estimada para CMC se asume igual al obtenido mediante RVR .

La definición de K_{CMC} permite redefinir el coeficiente $W_{CMC/RVR}$ como $= \frac{Q_{K,D}(G)(1-Q_{K,D}(G))}{K_{CMC}t_{RVR}\hat{V}_{RVR}}$.

5.1.4. Resultados Empíricos

A continuación se presentan los resultados numéricos obtenidos para los casos de estudio detallados en la Subsección 5.1.2.

Las estimaciones se obtienen para el mismo número de iteraciones $N = 10^6$ según la metodología mencionada.

5.1.4.1. Grafo completo K_n

El Cuadro 5.2 presenta los resultados numéricos para el grafo completo con n nodos, $n \in \{4, 5, 6, 7, 8\}$. Para cada valor de n se consideran los siguientes valores de D , $D = \{n - 1, \lfloor \frac{n}{2} \rfloor, 2\}$, y $q \in \{10^{-1}, 10^{-2}\}$.

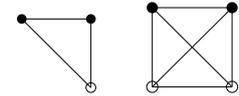


Figura 5.4: Grafo Completo, $n = 3, 4$

De acuerdo con los resultados presentados en el Cuadro 5.2, para un mismo valor de n , la estimación de la medida $Q_{K,D}(K_n)$ aumenta cuando D es menor. Este resultado es coherente dado que una cota más restrictiva disminuye los estados operativos posibles.

Por otro lado la varianza estimada para el mismo valor de n y q aumenta cuando la cota D es más pequeña. Lo mismo ocurre con el error relativo ER .

El valor de ER para un mismo n y D no evoluciona en un mismo sentido cuando disminuye q . Por ejemplo para $n = 4$ y $D = 2$ ER crece cuando q disminuye mientras que para $n = 7$ y $D = 3$, ER disminuye con q .

Cuadro 5.2: Resultados experimentales para el grafo completo K_n , $n = \{4, 5, 6, 7, 8\}$, con nodos objetivo $K = \{1, n\}$.

n	D	q	\hat{Q}_{RVR}	\hat{V}_{RVR}	ER	$V_{CMC/RVR}$	$W_{CMC/RVR}$
4	4	10^{-1}	2.1520×10^{-3}	1.602×10^{-12}	5.882×10^{-4}	1.340×10^3	7.244×10^7
4	4	10^{-2}	2.0230×10^{-6}	2.309×10^{-18}	7.512×10^{-4}	8.759×10^5	5.689×10^{10}
4	2	10^{-1}	3.6323×10^{-3}	8.497×10^{-10}	8.025×10^{-3}	4.259×10^0	2.767×10^5
4	2	10^{-2}	4.8902×10^{-6}	2.008×10^{-12}	2.898×10^{-1}	2.434×10^0	4.727×10^4
5	5	10^{-1}	2.0512×10^{-4}	6.484×10^{-15}	$.925 \times 10^{-4}$	3.162×10^4	1.303×10^9
5	5	10^{-2}	2.0004×10^{-8}	1.062×10^{-24}	5.152×10^{-5}	1.882×10^{10}	7.184×10^{14}
5	2	10^{-1}	6.6717×10^{-4}	2.454×10^{-10}	2.348×10^{-2}	2.716×10^0	8.386×10^4
5	2	10^{-2}	7.0006×10^{-8}	4.090×10^{-16}	2.889×10^{-1}	1.711×10^2	9.801×10^5
6	6	10^{-1}	2.0079×10^{-5}	6.280×10^{-18}	1.248×10^{-4}	3.196×10^6	7.682×10^{10}
6	6	10^{-2}	$2.0000007 \times 10^{-10}$	5.148×10^{-32}	1.134×10^{-6}	3.884×10^{15}	1.165×10^{20}
6	3	10^{-1}	2.0468×10^{-5}	3.330×10^{-14}	8.915×10^{-3}	6.145×10^2	4.170×10^6
6	3	10^{-2}	$2.0000006 \times 10^{-10}$	2.090×10^{-32}	7.229×10^{-7}	9.565×10^{15}	8.290×10^{19}
6	2	10^{-1}	1.4254×10^{-4}	7.347×10^{-11}	6.013×10^{-2}	1.939×10^0	5.646×10^3
6	2	10^{-2}	6.0513×10^{-10}	4.081×10^{-20}	3.338×10^{-1}	1.482×10^4	9.496×10^6
7	7	10^{-1}	2.0009×10^{-6}	2.044×10^{-21}	2.259×10^{-5}	9.787×10^8	1.183×10^{12}
7	7	10^{-2}	2.0303×10^{-12}	5.151×10^{-40}	1.134×10^{-8}	3.882×10^{21}	7.965×10^{25}
7	3	10^{-1}	2.0012×10^{-6}	3.225×10^{-20}	8.973×10^{-5}	6.205×10^7	6.471×10^{10}
7	3	10^{-2}	2.0000×10^{-12}	4.127×10^{-40}	1.015×10^{-8}	4.845×10^{21}	8.807×10^{25}
7	2	10^{-1}	2.3052×10^{-5}	1.069×10^{-11}	1.418×10^{-1}	2.155×10^0	3.004×10^2
7	2	10^{-2}	1.0203×10^{-10}	1.000×10^{-20}	9.802×10^{-1}	1.020×10^4	2.633×10^7
8	8	10^{-1}	2.0001×10^{-7}	1.720×10^{-25}	2.074×10^{-6}	1.162×10^{12}	2.769×10^{15}
8	8	10^{-2}	2.0000×10^{-14}	9.416×10^{-49}	4.852×10^{-11}	2.123×10^{28}	3.058×10^{32}
8	4	10^{-1}	2.0001×10^{-7}	1.197×10^{-24}	5.471×10^{-6}	1.670×10^{11}	2.781×10^{14}
8	4	10^{-2}	2.0000×10^{-14}	4.250×10^{-48}	1.030×10^{-10}	4.705×10^{27}	1.275×10^{32}
8	2	10^{-1}	4.4160×10^{-6}	2.126×10^{-12}	3.301×10^{-1}	2.077×10^0	6.681×10^1
8	2	10^{-2}	2.0000×10^{-14}	1.020×10^{-40}	5.050×10^{-7}	1.960×10^{20}	7.718×10^{23}

La estimación de $Q_{K,D}(K_n)$ disminuye cuando aumenta n lo que indica que el grafo completo es más confiable cuanto mayor es su tamaño. Se cumple $\ln \hat{Q}_{RVR} \propto (n-1) \ln q$. El valor de ER también disminuye cuando n es mayor, lo que indica mayor precisión en los grafos más confiables.

Para un mismo valor de q , la reducción de la varianza de RVR respecto de RVR es mayor en los grafos de mayor tamaño.

5.1.4.2. Grafo Grilla

El grafo Grilla se considera para 3 casos según $n \in \{3, 5, 6\}$ y conjunto de nodos objetivo $K = \{v \in V; \text{grado}(v) = 2\}$, $|K| = 4$. La Figura 5.5 presenta la grilla para $n = 3$. El Cuadro 5.3 detalla los resultados para distintos escenarios definidos según el valor de q , D y n .

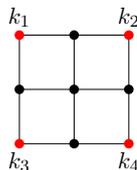


Figura 5.5: Red Grilla $n \times n$, $K = \{k_1, k_2, k_3, k_4\}$

Cuadro 5.3: Resultados experimentales para el grafo Grilla $n \times n$ con $K = \{v \in V : \text{grado}(v) = 2\}$.

n	D	q	\hat{Q}_{RVR}	\hat{V}_{RVR}	ER	$V_{CMC/RVR}$	$W_{CMC/RVR}$
3	8	10^{-3}	4.016857×10^{-6}	8.040×10^{-18}	7.059×10^{-4}	5.00×10^5	1.98×10^{10}
		10^{-6}	4.000009×10^{-12}	1.299×10^{-35}	9.013×10^{-7}	3.08×10^{17}	4.35×10^{20}
3	5	10^{-3}	1.401294×10^{-5}	9.000×10^{-12}	2.140×10^{-1}	1.56×10^0	7.05×10^4
		10^{-6}	1.000002×10^{-11}	5.999×10^{-24}	2.449×10^{-1}	1.67×10^6	2.83×10^{10}
5	24	10^{-3}	4.008037×10^{-6}	4.026×10^{-21}	1.583×10^{-5}	9.96×10^8	3.35×10^{10}
		10^{-6}	4.000006×10^{-12}	2.000×10^{-36}	3.536×10^{-7}	2.00×10^{18}	1.36×10^{21}
5	16	10^{-3}	4.007945×10^{-6}	3.941×10^{-21}	1.566×10^{-5}	1.02×10^9	3.16×10^{12}
		10^{-6}	4.000011×10^{-12}	7.000×10^{-36}	6.614×10^{-7}	5.71×10^{17}	4.60×10^{20}
5	8	10^{-3}	4.009004×10^{-6}	1.006×10^{-18}	2.502×10^{-4}	3.99×10^6	1.85×10^{11}
		10^{-6}	4.000009×10^{-12}	5.000×10^{-36}	5.590×10^{-7}	8.00×10^{17}	8.98×10^{19}
6	35	10^{-3}	4.007856×10^{-6}	3.858×10^{-21}	1.549×10^{-5}	1.04×10^9	6.28×10^{10}
		10^{-6}	4.000010×10^{-12}	5.999×10^{-36}	6.123×10^{-7}	6.67×10^{17}	2.77×10^{20}
6	29	10^{-3}	4.008055×10^{-6}	4.041×10^{-21}	1.586×10^{-5}	9.92×10^8	5.79×10^{10}
		10^{-6}	4.000008×10^{-12}	3.999×10^{-36}	4.999×10^{-7}	1.00×10^{18}	4.84×10^{20}
6	22	10^{-3}	4.008012×10^{-6}	4.028×10^{-21}	1.583×10^{-5}	9.95×10^8	6.07×10^{10}
		10^{-6}	4.000012×10^{-12}	7.999×10^{-36}	7.071×10^{-7}	5.00×10^{17}	2.46×10^{20}
6	15	10^{-3}	4.007825×10^{-6}	3.833×10^{-21}	1.544×10^{-5}	1.05×10^9	5.33×10^{10}
		10^{-6}	4.000010×10^{-12}	5.999×10^{-36}	6.123×10^{-7}	6.67×10^{17}	2.19×10^{20}

De acuerdo con los resultados detallados en el Cuadro 5.3, el orden de \hat{Q}_{RVR} es aproximadamente proporcional a q^2 . Para las distintas dimensiones de Grilla el error relativo oscila

entre 2.140×10^{-1} (para $n = 3$, $D = 5$ y $q = 10^{-3}$) y 9.013×10^{-7} (para $n = 3$, $D = 8$ y $q = 10^{-6}$).

RVR reduce la varianza en hasta de 2.00×10^{18} veces respecto de *CMC* (para $n = 5$, $D = 24$ y $q = 10^{-6}$) y su eficiencia supera a *CMC* en hasta 1.36×10^{21} veces (para $n = 5$, $D = 24$ y $q = 10^{-6}$).

5.1.4.3. Grafo Arpanet

El Cuadro 5.4 presenta los resultados para el grafo Arpanet (Figura 5.6), equiconfiable con nodos objetivo s y t . La anticonfiabilidad estimada \hat{Q}_{RVR} de Arpanet es aproximadamente proporcional a q^2 . Esto es coherente, pues la red es menos vulnerable a fallas cuanto más pequeño es q . Esta relación en general no se modifica cuando se reduce la cota de diámetro aunque aumenta sensiblemente la probabilidad de falla de la red. Por ejemplo, cuando $q = 10^{-1}$ y $D = 20$, $\hat{Q}_{RVR} = 9.5575 \times 10^{-2}$ mientras que para $q = 10^{-1}$ y $D = 10$, $\hat{Q}_{RVR} = 9.5594 \times 10^{-2}$. La evolución del error relativo se mantiene variando entre 1.3711×10^{-03} (para $D = 20$ y $q = 10^{-1}$) y 7.9999×10^{-01} (para $D = 10$ y $q = 10^{-6}$) para los distintos valores de D .

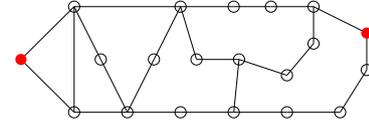


Figura 5.6: Arpanet

Cuadro 5.4: Resultados experimentales para el grafo Arpanet con nodos objetivo $K = \{s, t\}$.

D	q	\hat{Q}_{RVR}	\hat{V}_{RVR}	ER	$V_{CMC/RVR}$	$W_{CMC/RVR}$
20	10^{-1}	9.5575×10^{-2}	1.7172×10^{-8}	1.3711×10^{-3}	5.03×10^0	5.90×10^2
	10^{-2}	6.5568×10^{-4}	1.6903×10^{-11}	6.2704×10^{-3}	3.65×10^1	2.75×10^4
	10^{-3}	5.8964×10^{-6}	1.5459×10^{-14}	2.1087×10^{-2}	3.55×10^2	9.94×10^5
	10^{-4}	6.0814×10^{-8}	1.6314×10^{-17}	6.6417×10^{-2}	3.35×10^3	3.00×10^7
	10^{-5}	4.6001×10^{-9}	9.5996×10^{-21}	1.7625×10^{-1}	3.34×10^3	4.29×10^8
	10^{-6}	5.000002×10^{-12}	1.5999×10^{-23}	7.9999×10^{-1}	3.12×10^5	5.11×10^9
10	10^{-1}	9.5594×10^{-2}	1.7203×10^{-8}	1.3721×10^{-3}	4.81×10^0	5.61×10^2
	10^{-2}	6.5642×10^{-4}	1.6779×10^{-11}	6.2786×10^{-3}	3.67×10^1	2.91×10^4
	10^{-3}	6.1868×10^{-6}	1.6625×10^{-14}	2.0841×10^{-2}	3.52×10^2	8.38×10^5
	10^{-4}	6.1315×10^{-8}	1.6484×10^{-17}	6.6216×10^{-2}	3.56×10^3	3.22×10^7
	10^{-5}	6.0001×10^{-10}	2.1499×10^{-20}	2.1081×10^{-1}	3.21×10^4	2.88×10^8
	10^{-6}	5.000025×10^{-12}	1.5999×10^{-23}	7.9999×10^{-1}	3.13×10^5	3.33×10^9
6	10^{-1}	1.0136×10^{-1}	2.2663×10^{-8}	1.4852×10^{-3}	4.82×10^0	5.59×10^2
	10^{-2}	6.6588×10^{-4}	2.3254×10^{-11}	7.2419×10^{-3}	3.67×10^1	2.67×10^4
	10^{-3}	5.8991×10^{-6}	1.5482×10^{-14}	2.1093×10^{-2}	3.59×10^2	9.72×10^5
	10^{-4}	6.8415×10^{-8}	1.5355×10^{-17}	6.7081×10^{-2}	3.78×10^3	3.39×10^7
	10^{-5}	7.6001×10^{-10}	2.2399×10^{-20}	1.9692×10^{-1}	3.27×10^4	2.93×10^8
	10^{-6}	6.0000×10^{-12}	1.8000×10^{-23}	2.1213×10^{-1}	3.12×10^5	3.03×10^9

A diferencia de lo observado en la grilla el error relativo para un mismo valor de D aumenta cuando disminuye q , lo que indica pérdida de precisión en la estimación cuando la se vuelve más confiable.

Por otro lado el parámetro $W_{CMC/RVR}$ crece cuando disminuye q , alcanzando un valor de

5.11×10^9 cuando $q = 10^{-6}$ y $D = 20$. La mayor eficiencia de RVR respecto de CMC se registra cuando q es más pequeño, resultado esperado, pues RVR es más eficiente en redes altamente confiables.

5.1.4.4. Grafo Dodecahedro

La medida de anticonfiabilidad del grafo Dodecahedro se evalúa en dos casos: grafo equiconfiables con anticonfiabilidad q , y grafo con anticonfiabilidad elemental variando en dos valores, q_h y q_r , como se detalla en la Figura 5.8.

El Cuadro 5.5 presenta los resultados para el caso equiconfiable. De acuerdo a los valores obtenidos, el orden de \hat{Q}_{RVR} es aproximadamente proporcional a q^3 .

La reducción de varianza y el desempeño global del método RVR es superior en redes más confiables, con magnitudes elevadas en lo que respecta al valor de $V_{CMC/RVR}$. Éste alcanza valores de 5.00×10^{41} para las distintas magnitudes de D , lo que implica que RVR reduce la varianza de CMC en hasta 5.00×10^{41} .

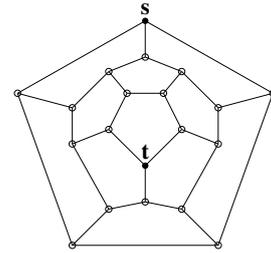


Figura 5.7: Dodecahedro equiconfiable

Por otra parte el error relativo aumenta cuando q disminuye. Para el caso $D = 19$, cuando $q = 10^{-1}$, $ER = 1.37 \times 10^{-3}$ y cuando $q = 10^{-6}$, $ER = 8 \times 10^{-1}$. En términos porcentuales el error varía entre 0.14 % y 80 %.

Cuadro 5.5: Resultados experimentales para el grafo Dodecahedro equiconfiable con $K = \{s, t\}$.

D	q	\hat{Q}_{RVR}	\hat{V}_{RVR}	ER	$V_{CMC/RVR}$	$W_{CMC/RVR}$
19	10^{-1}	2.88×10^{-3}	7.82×10^{-12}	9.70×10^{-2}	3.69×10^8	2.58×10^9
	10^{-2}	2.06×10^{-6}	2.33×10^{-18}	7.39×10^{-2}	1.32×10^{12}	2.02×10^{14}
	10^{-3}	2.01×10^{-9}	2.01×10^{-24}	7.06×10^{-2}	5.00×10^{17}	2.81×10^{20}
	10^{-4}	2.00×10^{-12}	2.02×10^{-34}	7.10×10^{-4}	4.95×10^{27}	7.66×10^{30}
	10^{-5}	2.00×10^{-15}	2.20×10^{-41}	2.35×10^{-4}	4.55×10^{34}	7.06×10^{37}
	10^{-6}	2.00×10^{-18}	2.00×10^{-48}	7.07×10^{-5}	5.00×10^{41}	7.81×10^{44}
10	10^{-1}	2.90×10^{-3}	1.90×10^{-11}	1.51×10^{-1}	1.52×10^8	8.50×10^8
	10^{-2}	2.06×10^{-6}	2.02×10^{-18}	6.90×10^{-2}	1.51×10^{12}	2.18×10^{14}
	10^{-3}	2.01×10^{-9}	1.00×10^{-24}	5.00×10^{-2}	9.98×10^{17}	4.77×10^{20}
	10^{-4}	2.01×10^{-12}	1.00×10^{-28}	4.97×10^{-1}	1.00×10^{22}	1.30×10^{25}
	10^{-5}	2.00×10^{-15}	2.20×10^{-41}	2.35×10^{-4}	4.55×10^{34}	7.09×10^{37}
	10^{-6}	2.00×10^{-18}	2.00×10^{-48}	7.07×10^{-5}	5.00×10^{41}	1.32×10^{45}
5	10^{-1}	1.24×10^{-2}	9.37×10^{-9}	7.81×10^{-1}	1.32×10^6	9.09×10^5
	10^{-2}	4.06×10^{-6}	2.00×10^{-12}	3.48×10^1	2.53×10^6	7.48×10^7
	10^{-3}	2.00×10^{-9}	2.11×10^{-27}	2.29×10^{-3}	4.75×10^{20}	1.12×10^{23}
	10^{-4}	2.00×10^{-12}	2.00×10^{-34}	7.07×10^{-4}	5.00×10^{27}	3.75×10^{30}
	10^{-5}	2.00×10^{-15}	1.30×10^{-41}	1.80×10^{-4}	7.69×10^{34}	6.88×10^{37}
	10^{-6}	2.00×10^{-18}	problemas numéricos (varianza negativa)			

El grafo Dodecahedro con q variable en dos valores, q_h y q_r se asigna según indica la Figura 5.8. Los resultados numéricos se presentan en el Cuadro 5.6.

La configuración de los valores q_h y q_r modifica la probabilidad de falla de la red \hat{Q}_{RVR} . El valor de la anticonfiabilidad cuando q varía en $\{q_h, q_r\}$ registra un valor intermedio a los valores que registra en una red equiconfiable con q igual a q_r o q_h . Si:

- Q_i es la anticonfiabilidad de la red cuando los enlaces son equiconfiables $q_i, i \in \{r, h\}$,
- $Q_{r,h}$ es la anticonfiabilidad de la red cuando q diferenciado en dos valores, $\{q_h, q_r\}$,

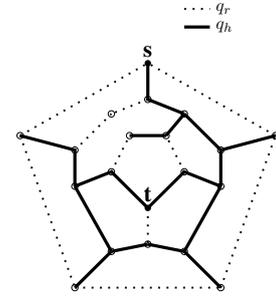


Figura 5.8: Dodecahedro, para q variando en 2 valores, $\{q_h, q_r\}$.

y se cumple $q_r < q_h$ entonces:

$$Q_r \leq Q_{r,h} \leq Q_h.$$

Por ejemplo la medida \hat{Q}_{RVR} , para $D = 19$ y enlaces equiconfiables con $q_h = 10^{-3}$ es $Q = 2.01 \times 10^{-9}$, mientras que si $q_r = 10^{-5}$ la anticonfiabilidad es 2.00×10^{-15} . El dodecahedro con q variable en $\{10^{-3}, 10^{-5}\}$ tiene anticonfiabilidad 2.1044×10^{-11} .

El parámetro $W_{CMC/RVR}$ varía entre 2.187×10^7 y 1.911×10^{27} .

Cuadro 5.6: Resultados experimentales para el grafo Dodecahedro con q variable en 2 valores $\{q_h, q_r\}$, $K = \{s, t\}$.

D	(q_h, q_r)	\hat{Q}_{RVR}	\hat{V}_{RVR}	ER	$V_{CMC/RVR}$	$W_{CMC/RVR}$
19	$(10^{-6}, 10^{-5})$	2.0000×10^{-17}	9.99×10^{-49}	4.999×10^{-6}	2.000×10^{25}	8.834×10^{21}
	$(10^{-3}, 10^{-5})$	2.1044×10^{-11}	1.026×10^{-27}	1.522×10^{-1}	2.050×10^{10}	2.190×10^7
	$(10^{-4}, 10^{-3})$	2.001×10^{-11}	8.719×10^{-33}	4.666×10^{-4}	2.269×10^{15}	1.492×10^{12}
	$(10^{-3}, 10^{-4})$	2.0118×10^{-10}	9.840×10^{-28}	1.559×10^{-2}	2.044×10^{11}	2.914×10^8
10	$(10^{-6}, 10^{-5})$	2.0000×10^{-17}	3.250×10^{-54}	9.013×10^{-9}	6.154×10^{30}	1.911×10^{27}
	$(10^{-3}, 10^{-5})$	2.1002×10^{-11}	9.821×10^{-28}	1.492×10^{-1}	2.138×10^{10}	2.187×10^7
	$(10^{-4}, 10^{-3})$	2.0012×10^{-11}	1.061×10^{-32}	5.146×10^{-4}	1.86×10^{15}	1.036×10^{12}
	$(10^{-3}, 10^{-4})$	2.0122×10^{-10}	1.033×10^{-27}	1.598×10^{-2}	1.977×10^{11}	2.229×10^8
5	$(10^{-6}, 10^{-5})$	2.0000×10^{-17}	1.999×10^{-48}	7.070×10^{-6}	1.000×10^{25}	6.097×10^{21}
	$(10^{-3}, 10^{-5})$	2.1030×10^{-11}	1.012×10^{-27}	1.512×10^{-1}	2.077×10^{10}	5.989×10^7
	$(10^{-4}, 10^{-3})$	2.2012×10^{-11}	9.690×10^{-33}	4.920×10^{-4}	1.868×10^{15}	1.036×10^{12}
	$(10^{-3}, 10^{-4})$	2.0123×10^{-10}	1.034×10^{-27}	1.598×10^{-2}	1.926×10^{11}	7.450×10^8

5.1.4.5. Grafo Red ANTEL

El análisis de la red de transporte de ANTEL (Figura 5.9) se basa en dos subredes de la red original, identificadas según región este y oeste como muestra la Figura 5.10.

En ambos casos se consideran enlaces equiconfiables y conjunto de nodos objetivo $K = \{v \in V; grado(v) \geq 3\}$.

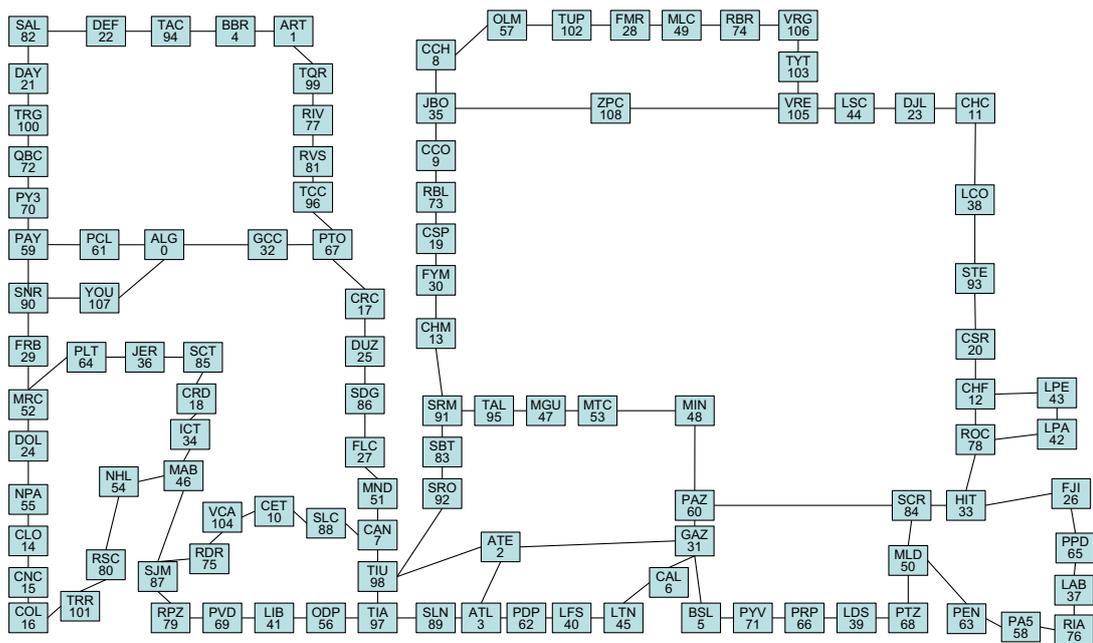


Figura 5.9: Red de transporte de ANTEL

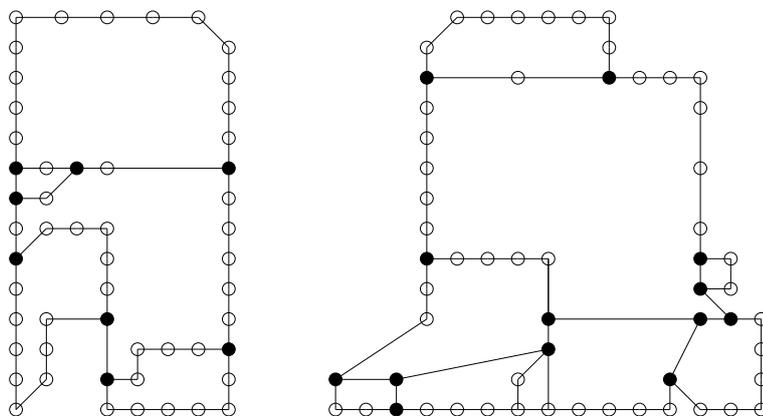


Figura 5.10: Red de transporte ANTEL: región oeste y este respectivamente

Los Cuadros 5.7 y 5.8 detallan los resultados para la región este y oeste respectivamente. En todos los casos de prueba la anticonfiabilidad estimada de la región este es superior a la registrada en la región oeste, lo que indica mayor vulnerabilidad a fallas en la región este. El error relativo es elevado. Los valores obtenidos varían entre 9.76×10^{-3} (para región este, $D = 58$, $q = 10^{-3}$) y 5.86×10^{-1} (para región oeste, $D = 40$ y $q = 10^{-6}$).

Cuadro 5.7: Resultados experimentales para la red ANTEL, zona oeste, conjunto de nodos objetivo $K = \{v \in V; \text{grado}(v) \geq 3\}$.

D	q	\hat{Q}_{RVR}	\hat{V}_{RVR}	ER
53	10^{-3}	2.1029×10^{-5}	6.4976×10^{-14}	1.21×10^{-2}
	10^{-6}	1.6000×10^{-11}	4.4001×10^{-23}	4.15×10^{-1}
40	10^{-3}	2.1199×10^{-5}	6.5745×10^{-14}	1.21×10^{-2}
	10^{-6}	8.0002×10^{-12}	2.2000×10^{-23}	5.86×10^{-1}
30	10^{-3}	2.6511×10^{-5}	6.0624×10^{-12}	9.29×10^{-2}
	10^{-6}	1.3000×10^{-11}	3.5001×10^{-23}	4.55×10^{-1}

Cuadro 5.8: Resultados experimentales para la red ANTEL, zona este, conjunto de nodos objetivo $K = \{v \in V; \text{grado}(v) \geq 3\}$.

D	q	\hat{Q}_{RVR}	\hat{V}_{RVR}	ER
58	10^{-3}	5.6361×10^{-5}	3.0241×10^{-13}	9.76×10^{-3}
	10^{-6}	3.4000×10^{-11}	1.4700×10^{-22}	3.57×10^{-1}
40	10^{-3}	5.6560×10^{-5}	3.0402×10^{-13}	1.04×10^{-2}
	10^{-6}	3.4000×10^{-11}	3.4299×10^{-22}	2.98×10^{-1}
30	10^{-3}	5.5421×10^{-5}	2.9576×10^{-13}	1.21×10^{-2}
	10^{-6}	5.5000×10^{-11}	2.9399×10^{-22}	3.11×10^{-1}

5.1.5. Comentarios generales

El método *RVR* es más eficiente que *CMC* en todos los casos de prueba elegidos, según indican los valores de \hat{V}_{RVR} y $W_{CMC/RVR}$.

Los valores obtenidos para ER , $W_{CMC/RVR}$ y $V_{CMC/RVR}$ dan cuenta de la heterogeneidad de los resultados que varían de acuerdo a la topología objeto de estudio.

En el Cuadro 5.9 se presentan los valores extremos para los tres indicadores. En todos los casos de prueba el método *RVR* reduce la varianza y es más eficiente que el método *CMC*.

Cuadro 5.9: Mínimo y máximo registrado en todos los casos de prueba para ER , $W_{CMC/RVR}$ y $V_{CMC/RVR}$.

indicador	mín	máx
ER	4.85×10^{-11}	9.80×10^{-1}
$V_{CMC/RVR}$	1.56×10^0	5.00×10^{41}
$W_{CMC/RVR}$	3.00×10^2	1.32×10^{45}

La relación $V_{CMC/RVR}$ y la evolución de ER cuando q disminuye varía según los casos de prueba. En los grafos K_6 , K_7 , K_8 y la Grilla $n \times n$, el valor de ER disminuye cuando q disminuye. En cambio, en Arpanet y en las subredes de ANTEL el valor de ER aumenta cuando q disminuye.

5.2. Diagnóstico empírico

En esta sección se presenta el diagnóstico de robustez del método RVR en base a la función de cobertura definida por Schruben [57].

Se consideran tres topologías: el grafo puente, el grafo dodecahedro y la red ANTEL, región oeste. En las dos primeras se evalúa el estimador $Q_{\{s,t\},D}(G)$ y en el tercero $Q_{K,D}(G)$. En todos los casos se asumen enlaces equiconfiables.

Los resultados indican que el error relativo evoluciona de manera diferenciada cuando disminuye q . En algunas topologías el error relativo es decreciente con q y en otras creciente. La dependencia de la calidad de las estimaciones respecto de la magnitud de q evidencia que en el caso general el método no es robusto.

5.2.1. Función de Cobertura

La función de cobertura es una medida de robustez que permite analizar empíricamente la validez de estimadores por intervalo.

Si se define un estimador \hat{R} de R y se obtiene el intervalo de confianza de cobertura $(1 - \alpha)$, IC_α , entonces se asume que la probabilidad de que el valor verdadero pertenezca al intervalo es $1 - \alpha$:

$$\Pr(R \in IC_\alpha) = 1 - \alpha.$$

Empíricamente, de cada M intervalos que se estiman, $(1 - \alpha) \times M$ contienen el verdadero valor R :

$$F_{\nu^*}(\nu) = \frac{1}{M} \sum_{i=1}^M 1_{\{\nu^* \leq \nu\}}, \quad (5.1)$$

con $F_{\nu^*}(\nu)$ la distribución empírica de ν^* . Cada intervalo es aleatorio y se define como una función de los vectores de estado de la red.

En el caso de confiabilidad, si el método es robusto entonces para cualquier coeficiente de cobertura α y para cualquier probabilidad de falla q el intervalo de confianza estimado debe asegurar cobertura $1 - \alpha$. Si esto es cierto la función de cobertura se distribuye uniformemente

en el intervalo $[0, 1]$.

La definición del intervalo de confianza se basa en el Teorema del Límite Central 5.2.1,

Teorema 5.2.1 *Si las variables X_1, X_2, \dots, X_n son independientes, equidistribuidas, con esperanza μ y varianza σ^2 entonces para cada x se cumple:*

$$\lim_{n \rightarrow \infty} \text{Prob} \left(\frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n\sigma^2}} \leq x \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$$

La expresión $\frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n\sigma^2}}$ converge en distribución a la Normal típica.

En el modelo de confiabilidad, el parámetro de interés es el promedio de los N configuraciones sorteadas, $Q_{K,D}(G)$ y como se cumple $\hat{V} < \infty$ puede definirse el intervalo según:

$$IC_\alpha = \left[\hat{Q}_{K,D}(G) - z_{\frac{\alpha}{2}} \hat{V}^{\frac{1}{2}}, \hat{Q}_{K,D}(G) + z_{\frac{\alpha}{2}} \hat{V}^{\frac{1}{2}} \right],$$

donde $z_{\frac{\alpha}{2}} = \Phi^{-1}(1 - \frac{\alpha}{2})$ y $\hat{Q}_{K,D}(G)$ el estimador de la medida. La propiedad $\hat{V} < \infty$ se prueba trivialmente.

Demostración. El método *RVR* asegura reducción de varianza respecto de *CMC* [39] y la medida de anticonfiabilidad $\hat{Q}_{K,D} \in [0, 1]$. Se prueba entonces que \hat{V}_{RVR} está acotado superiormente:

$$\hat{V}_{RVR} \leq \hat{V}_{CMC} = \hat{Q}_{K,D}(1 - \hat{Q}_{K,D}) \leq \frac{1}{4}.$$

QED

Si el estimador es robusto la función de distribución empírica 5.1 converge a la distribución Uniforme(0,1), $F(\nu) = \nu$, para cualquier topología y cualquier valor de anticonfiabilidad elemental q .

En resumen, el diagnóstico evalúa la distancia entre la función de cobertura estimada y la función de distribución $F(x) = x$. Si el estimador es robusto entonces $|F_{\nu^*}(\nu) - F(\nu)| < \epsilon$, $\forall \nu$. La comparación entre ambas funciones se basa en la representación gráfica, que proporciona de manera inmediata una noción de distancia entre la función empírica y la teórica, y se completa con una prueba de bondad de ajuste no paramétrica de Kolmogorov-Smirnov [58].

5.2.1.1. Casos de prueba

La función de cobertura se obtiene para el grafo Puente, Dodecahedro y Antel, región oeste (Ver Figura 5.11).

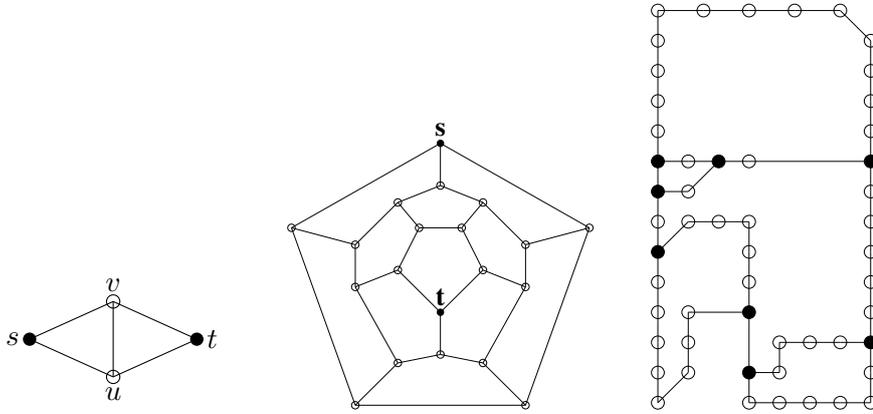


Figura 5.11: Casos elegidos para evaluar la función de cobertura

La estimación de la función de cobertura se basa en los siguientes parámetros:

- $\alpha = \{0, 0.005, 0.01, 0.0015, \dots, 1\}$,
- $N = 10^6$,
- $D = n - 1$
- el cálculo exacto de la medida $Q_{K,D}(G)$ se obtiene mediante el Algoritmo 14, derivado de RVR,
- se realizan 200 réplicas de la medida de $Q_{K,D}(G)$ para estimar la función de distribución empírica $F_{\nu^*}(\nu)$.
- se consideran valores de $q \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}\}$.

5.2.2. Resultados

Los casos presentados en la Subsección 5.2.1.1 evidencian cómo la función de distribución empírica $F_{\nu^*}(\nu)$ se aleja de la distribución Uniforme, alejándose de la cobertura deseada cuando disminuye la anticonfiabilidad elemental q .

Las Figuras 5.12, 5.13 y 5.14 representa la función de cobertura para el grafo puente, dodecahedro y ANTEL región oeste respectivamente.

La prueba de hipótesis de Kolmogorov Smirnov con nivel de confianza 5 % permite formalizar esta idea.

El Cuadro 5.10 presenta los datos para el grafo puente. Para valores de q superior a 10^{-2} se rechaza la hipótesis nula.

Los resultados para el dodecahedro se detallan en el Cuadro 5.11. A partir de $q = 10^{-4}$ se rechaza la hipótesis de que la función empírica se distribuye uniformemente en $(0,1)$.

Por último el Cuadro 5.12 resume los resultados para la red ANTEL. Para valores de q entre 6×10^{-1} y 10^{-3} no se rechaza que la distribución de cobertura empírica es Uniforme(0,1). No ocurre lo mismo para $q \in \{10^{-4}, 10^{-5}, 10^{-6}\}$.

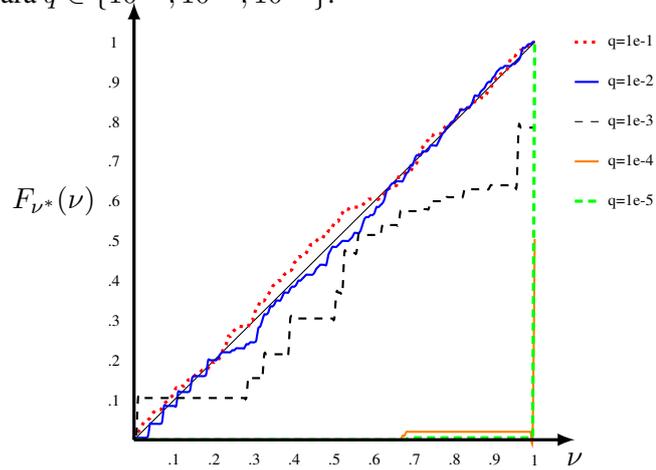


Figura 5.12: Función de cobertura para grafo Puente, $Q_{\{s,t\},(n-1)}(G)$ y $q \in \{10^{-1}, 10^{-6}\}$

Cuadro 5.10: Prueba de Kolmogorov-Smirnov para grafo puente.

q	Rechazo H_o	p-valor
10^{-5}	SI	1.21×10^{-174}
10^{-4}	SI	2.09×10^{-169}
10^{-3}	SI	4.35×10^{-18}
10^{-2}	NO	0.4859
10^{-1}	NO	0.5053

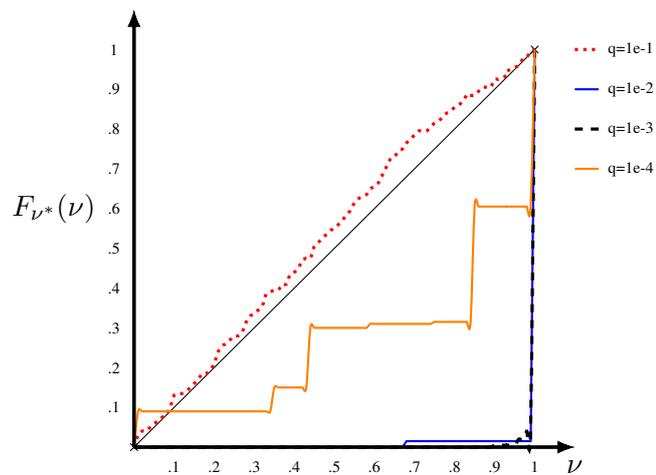
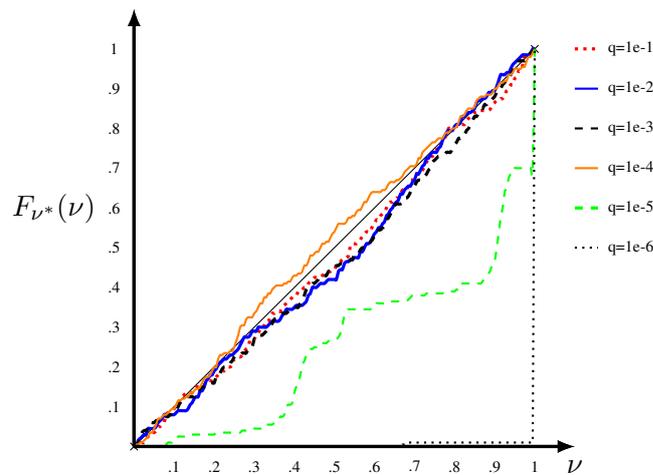


Figura 5.13: Función de cobertura para grafo Dodecahedro, $Q_{\{s,t\},(n-1)}(G)$ y $q \in \{10^{-1}, 10^{-6}\}$

Cuadro 5.11: Prueba de Kolmogorov-Smirnov para grafo Dodecahedro.

q	Rechazo H_o	p-valor
10^{-6}	SI	2.17×10^{-25}
10^{-5}	SI	9.99×10^{-82}
10^{-4}	SI	5.02×10^{-86}
10^{-3}	NO	0.343

Figura 5.14: Función de cobertura: para red de ANTEL, región oeste, $Q_{\{s,t\},(n-1)}(G)$ y $q \in \{10^{-1}, 10^{-6}\}$

Cuadro 5.12: Prueba de Kolmogorov-Smirnov para red de ANTEL, región oeste.

q	Rechazo H_o	p-valor
10^{-6}	SI	3.63×10^{-90}
10^{-5}	SI	6.88×10^{-173}
10^{-4}	SI	2.47×10^{-37}
10^{-3}	NO	0.6245
10^{-2}	NO	0.1270
10^{-1}	NO	0.0874
2×10^{-1}	NO	0.5082
3×10^{-1}	NO	0.9826
4×10^{-1}	NO	0.6647
5×10^{-1}	NO	0.0807
6×10^{-1}	NO	0.2397

De acuerdo a los resultados el método RVR en la versión implementada no es robusto incluso para valores de q que no son demasiado pequeños como para considerar el evento de falla como un evento raro.

5.3. Estudio asintótico del Error Relativo

En contexto de eventos raros, un método es robusto si el error relativo permanece acotado (Se dice que el método cumple *BRE*, Bounded Relative Error) cuando la anticonfiabilidad elemental tiende a 0.

Un método es *BRE* si sólo si [23]:

$$\lim_{q \rightarrow 0} \frac{E(\hat{Q}^2)}{E(\hat{Q})^2} \rightarrow 0. \quad (5.2)$$

A continuación se detalla el cálculo del límite para un grafo sencillo a partir del cual se prueba que el método aplicado no es robusto.

Proposición 5.3.1 *El método RVR en su versión multinomial y para un conjunto de corte arbitrario no cumple BRE.*

Demostración.

Sea $G = (V, E)$ el grafo de la Figura 5.15, equiconfiable, con nodos objetivo $K = \{s, t\}$, anticonfiabilidad q y $D = 2$. $Y_G = 1_{\{G \text{ falla}\}}$.

Sea C un D -cutset conformado por los enlaces $\{(t, u), (t, s)\}$ y G_1, G_2 los grafos obtenidos a partir de G . G_1 se obtiene fijando la arista (t, u) en estado operativo, y G_2 se obtiene de establecer en estado de falla (t, u) y fijar operativa (t, s) .

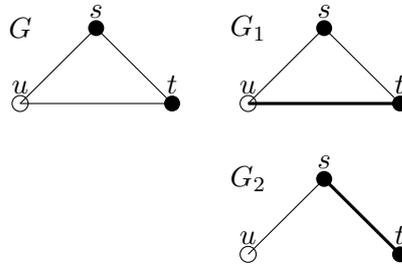


Figura 5.15: Grafo triángulo

El estimador insesgado $\frac{S1(G)}{N}$ de la medida de anticonfiabilidad se define como:

$$\frac{S1(G)}{N} = \frac{1}{N} \left(NQ_C + (1 - Q_C) \sum_{i=1}^{|C|} S1(G_i) \right).$$

Se cumple

- $Q_{\{s,t\},2}(G_1) > 0$
- $Q_{\{s,t\},2}(G_2) = 0$.

- $Q_C = q^2$, es la probabilidad de falla del corte.
- $P_G = \{A_F, A_1, A_2\}$, es la partición de Ω_G definida a partir de C :
 - A_F : conjunto de vectores de estado del grafo tal que los enlaces de C fallan.
 - A_1 : conjunto de vectores de estado del grafo tal que (t, u) opera.
 - A_2 : conjunto de vectores de estado del grafo tal que (t, u) falla y (t, s) opera.
- Y_{G_i} es el estado de G , condicionado al componente i -ésimo de la partición de Ω_G , A_i , $i \in \{1, 2\}$.
- $E(Y_{G_1}) = \Pr(Y_{G_1} = 1) = q^2$. Es la probabilidad de que el grafo G_1 falle dado que (t, u) está operativa.
- $E(Y_{G_2}) = \Pr(Y_{G_2} = 1) = 0$. Es la probabilidad de falla del grafo G_2 dado que (t, u) falla y (s, t) opera. En esta caso G_2 está operativo siempre.
- $E(Y_{G_1}^2) = \Pr(Y_{G_1} = 1) = q^2$
- $E(Y_{G_2}^2) = \Pr(Y_{G_2} = 1) = 0$
- Si N es un vector que contiene la muestra multinomial asociada al método RVR entonces $N = (N_1, N_2)$, con N_i variables aleatorias de distribución Binomial(p_i) donde:
 - $p_1 = \Pr(N_1 = n_1) = \frac{1-q}{1-q^2}$.
 - $p_2 = \Pr(N_2 = n_2) = \frac{rq}{1-q^2}$.
 - $E(N_1) = \frac{N(1-q)}{1-q^2}$.
 - $E(N_2) = \frac{N(1-q)q}{1-q^2}$.
 - $Var(N_1) = N \frac{rq}{1-q^2} (1 - \frac{rq}{1-q^2}) = \frac{Nq(1-q)^2}{(1-q^2)^2}$.
 - $E(N_1^2) = Var(N_1) + E(N_1)^2 = \frac{N(1-q)^2}{(1-q^2)^2} (q + N)$.
- $E(S1(G_1)) = q^2 \frac{r}{1-q^2} N$ y $E(S1(G_2)) = 0$.

Demostración.

$$\begin{aligned}
 E(S1(G_1)) &= E\left(\sum_{j=1}^{N_1} Y_{G_1}^j\right) = E\left(E\left(\sum_{j=1}^{N_1} Y_{G_1}^j / N_1\right)\right) = E\left(\sum_{j=1}^{N_1} E(Y_{G_1}^j)\right) = \\
 &= E(N_1 E(Y_{G_1})) = E(N_1 q^2) = q^2 E(N_1) = q^2 \frac{r}{1-q^2} N.
 \end{aligned}$$

En forma análoga, usando que $E(Y_{G_2}^2) = 0$, se deduce que $E(S1(G_2)) = 0$.

QED

- $E(S1(G_1)^2) = \frac{q^2 N(1-q)}{2(1-q^2)^2} (2 + q^3 - 3q^2)$ y $E(S1(G_2)^2) = 0$.

Demostración.

$$E(S1(G_1)^2) = E \left(\left(\sum_{j=1}^{N_1} Y_{G_1}^j \right)^2 \right) = E \left(E \left(\sum_{j=1}^{N_1} Y_{G_1}^j \right)^2 / N_1 \right)$$

$$E \left(E \left(\sum_{j=1}^{N_1} (Y_{G_1}^j)^2 + 2 \sum_{k < j}^{N_1} Y_{G_1}^j Y_{G_1}^k \right) \right) \quad (5.3)$$

Además:

- $E \left(\sum_{j=1}^{N_1} (Y_{G_1}^j)^2 \right) = N_1 E((Y_{G_1})^2) = N_1 q^2$,
- $E \left(\sum_{k < j}^{N_1} Y_{G_1}^j Y_{G_1}^k \right) = \frac{N_1(N_1-1)q^4}{2}$.

Sustituyendo en la Ecuación 5.3 se obtiene:

$$E(S1(G_1)^2) = E \left(N_1 q^2 + \frac{N_1(N_1-1)}{2} q^4 \right) = q^2 E(N_1) + \frac{q^4}{2} (E(N_1^2) - E(N_1)) =$$

$$E(S1(G_1)^2) = \frac{q^2 N(1-q)}{2(1-q^2)^2} (2 + q^3 - 3q^2) \quad (5.4)$$

QED

- $E(S1(G_1)S1(G_2)) = 0$ pues G_2 nunca falla.

La esperanza de $S1(G)$ para el grafo de la Figura 5.15 es:

$$E(S1(G)) = E(NQ_C + (1-Q_C) \sum_{i=1}^2 S1(G_i)) = Nq^2 + (1-q^2)(E(S1(G_1)) + E(S1(G_2)))$$

$$E(S1(G)) = Nq^2 + N(1-q^2) \left(q^2 \frac{1-q}{1-q^2} \right) = Nq^2(2-q)$$

$$E(S1(G)) = Nq^2(2-q) \quad (5.5)$$

Por otro lado la esperanza del cuadrado $E(S1(G)^2)$:

$$E(S1(G)^2) = E \left(N^2 Q_C^2 + (1-Q_C)^2 \left(\sum_{i=1}^{|C|} S1(G_i) \right)^2 + 2NQ_C(1-Q_C) \sum_{i=1}^{|C|} S1(G_i) \right)$$

$$= E(N^2 q^4 + (1-q^2)^2 (S1(G_1) + S1(G_2))^2 + 2Nq^2(1-q^2)(S1(G_1) + S1(G_2)))$$

$$= N^2 q^4 + (1-q^2)^2 E(S1(G_1)^2 + S1(G_2)^2 + 2S1(G_1)S1(G_2)) + 2Nq^2(1-q^2)E(S1(G_1) + S1(G_2))$$

$$E(S1(G)^2) = N^2q^4 + \frac{q^2N(1-q)}{2}(2 + q^3 - 3q^2) + 2N^2q^4(1-q)$$

$$E(S1(G)^2) = Nq^2(Nq^2 + \frac{(1-q)}{2}(2 + q^3 - 3q^2) + 2Nq^2(1-q)). \quad (5.6)$$

De acuerdo con las Ecuaciones 5.5 y 5.6 el límite del cociente 5.3 cumple:

$$\lim_{q \rightarrow 0} \frac{E(S1(G)^2)}{E(S1(G))^2} = \lim_{q \rightarrow 0} \frac{Nq^2(Nq^2 + \frac{(1-q)}{2}(2 + q^3 - 3q^2) + 2Nq^2(1-q))}{N^2q^4(2-q)^2}$$

$$\simeq \lim_{q \rightarrow 0} \frac{Nq^2}{N^2q^4} = \lim_{q \rightarrow 0} \frac{1}{Nq^2} \rightarrow \infty.$$

El método RVR no cumple BRE.

QED

El gráfico de la Figura 5.16 muestra la evolución del error relativo para el grafo de la Figura 5.15 para algunos valores de q . El eje horizontal representa los valores de q en orden descendente y el eje de las ordenadas representa el error relativo correspondiente. El Cuadro 5.13 detalla los valores graficados.

De acuerdo a la simulación el error relativo para el triángulo crece cuando se reduce q y de acuerdo al cálculo del límite no está acotado superiormente.

Cuadro 5.13: Evolución de ER para el grafo de la Figura 5.15. Calculado en base a RVR versión multinomial.

q	ER
9.00×10^{-1}	9.58×10^{-5}
8.00×10^{-1}	1.86×10^{-4}
7.00×10^{-1}	2.76×10^{-4}
6.00×10^{-1}	3.69×10^{-4}
5.00×10^{-1}	4.71×10^{-4}
4.00×10^{-1}	5.93×10^{-4}
3.00×10^{-1}	7.52×10^{-4}
2.00×10^{-1}	9.94×10^{-4}
1.00×10^{-1}	1.50×10^{-3}
1.00×10^{-2}	4.97×10^{-3}
1.00×10^{-3}	1.58×10^{-2}
1.00×10^{-4}	4.99×10^{-2}
1.00×10^{-5}	1.57×10^{-1}
1.00×10^{-6}	4.71×10^{-1}

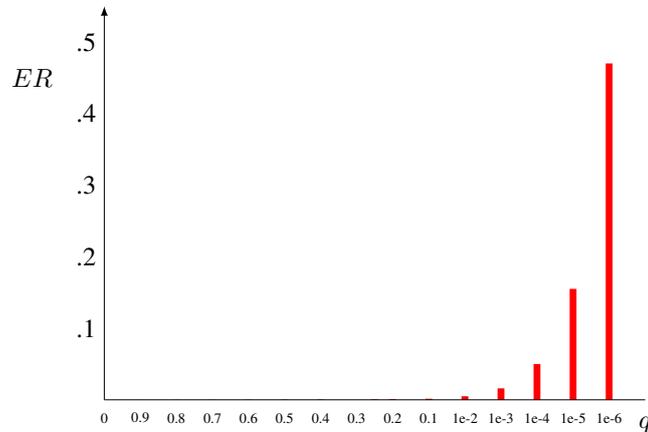


Figura 5.16: Evolución de ER para el grafo de la Figura 5.15. Calculado en base a RVR versión multinomial.

5.4. Conclusiones

De acuerdo a las estimaciones obtenidas la medida $Q_{K,D}(G)$ tiene relación inversa con la cota de diámetro: valores pequeños de D se traducen en una red menos confiable. Esta propiedad es previsible dado que se reduce la dimensión del espacio de vectores de estados que generan estados operativos de G .

Por otro lado el cociente $W_{CMC/RVR}$ indica que la eficiencia del método RVR es superior a CMC , en particular en redes altamente confiables.

Sin embargo, aunque RVR reduce tanto la varianza como el tiempo de cómputo en las distintas topologías consideradas, el diagnóstico evidencia que el método implementado no es robusto.

La calidad de las estimaciones se deteriora al disminuir la anticonfiabilidad elemental de los enlaces. De acuerdo con los resultados empíricos, el error relativo no evoluciona de la misma forma para distintas topologías cuando se reduce el valor de q . En la red ARPANET así como en las subredes derivadas de la red de ANTEL y el dodecahedro equiconfiable el error relativo aumenta cuando disminuye q mientras que en la Grilla el error disminuye.

El diagnóstico basado en la función de cobertura muestra como la cobertura empírica del intervalo de confianza se aleja de la cobertura deseada cuando q es pequeño. Esto evidencia la dependencia que de la cobertura empírica respecto de la magnitud de q .

Por otro lado la instancia simple (basada en un triángulo) demuestra que el método RVR no verifica la propiedad BRE.

En base a los resultados empíricos y teóricos se concluye que el método en la versión implementada no es adecuado para estimar la medida $Q_{K,D}(G)$ dado que la magnitud del error relativo depende de la magnitud de q .

Parte IV

CONCLUSIONES

Capítulo 6

Conclusiones Generales

La confiabilidad diámetro acotada en redes es un concepto definido recientemente y el desarrollo teórico y empírico al respecto es escaso si se lo compara con su predecesor, la medida de confiabilidad clásica. En este sentido son diversas las temáticas disponibles para aportar a esta nueva medida, que permite evaluar la capacidad de que un conjunto de nodos puedan comunicarse por caminos operativos de largo menor o igual a D .

El trabajo desarrollado en esta tesis estima la medida de anticonfiabilidad diámetro acotada con $K \subset V$ para distintas topologías, magnitudes de diámetro D y confiabilidad elemental q . Las estimaciones se basan en el método recursivo de reducción de varianza (*RVR*) y se generan en MATLAB, en el cluster¹ de Facultad Ingeniería.

El método *RVR* se basa en la identificación de una partición del espacio de estados del grafo, definida en función de un conjunto de corte, y aplica una relación recursiva con muestreo multinomial. De acuerdo con los resultados *RVR* tiene mejor desempeño que el método *CMC*, generando estimaciones de menor varianza en menos tiempo en todos los casos de estudio.

Además de caracterizar los valores que toma la medida $Q_{K,D}(G)$ en distintos escenarios se evalúa la robustez del método respecto de la magnitud de q .

En la actualidad las redes de comunicación son altamente confiables y los eventos de falla en los componentes ocurren con baja probabilidad. De esta forma, si bien el método *RVR* mejora el desempeño de *CMC* es relevante evaluar la calidad de las estimaciones.

El diagnóstico empírico, basado en estudio de casos y en la función de cobertura, brindan evidencia de la dependencia existente entre la calidad de las estimaciones obtenidas y la magnitud de q . El estudio de casos permite concluir que el error relativo no siempre se reduce al disminuir q , evidenciando la pérdida de precisión del método cuando el evento de falla se hace más raro. La función de cobertura empírica obtenida para tres topologías puntuales evidencia que la cobertura de los intervalos de confianza estimados se aleja de la cobertura deseada cuando disminuye q .

¹<http://www.fing.edu.uy/cluster>

Por otro lado el estudio asintótico del error relativo para un grafo sencillo prueba que la versión implementada de *RVR* no cumple *BRE*.

En resumen, la implementación elegida en esta tesis para estimar la medida de confiabilidad diámetro acotada indica que *RVR* tiene mayor precisión y eficiencia que *CMC* en todos los casos de prueba. Sin embargo la precisión en las estimaciones se ve afectada por la magnitud de la anticonfiabilidad elemental q y por lo tanto el método no es robusto.

6.1. Trabajo futuro

El desarrollo de este trabajo se enmarca en el Proyecto CSIC I+D “Propiedades de la confiabilidad diámetro-acotada de redes y métodos de cálculo”. El estudio de la diámetro confiabilidad y el desempeño del método *RVR* no culmina con esta tesis sino que es el inicio de una etapa de investigación enmarcada en un proyecto conjunto.

Algunas de las líneas en las que queda pendiente profundizar son :

- investigar y mejorar el proceso de selección del conjunto de corte en cada paso recursivo de *RVR*,
- incorporar mecanismos de reducción de grafos,
- evaluar la influencia de la forma de selección de corte en la robustez del método aplicado a la estimación de la arista-confiabilidad diámetro acotada,
- extender las aplicaciones al caso V –terminal.
- estudiar la posibilidad de adaptar las reducciones serie-paralelo para aplicarlas al estudio de la diámetro confiabilidad.

Bibliografía

- [1] M. O. Ball, C. J. Colbourn, and J. S. Provan, “Network reliability,” Institute for Systems Research Technical Reports, Tech. Rep. 92-74, 1992.
- [2] M. O. Locks and A. Satyarayana, “Network reliability,” in *The State of Art*, 1986.
- [3] C. J. Colbourn, *The Combinatorics of Network Reliability*. Oxford University Press, Inc New York, USA, 1987.
- [4] G. Rubino, “Network reliability evaluation,” in *State of art in performance modeling and simulation*, K. Bagchi and J. Walrand, Eds., 1998, pp. 275–302.
- [5] M. Ball, *Computational complexity of network reliability analysis: An overview*. IEEE Transactions on Reliability, 1986, vol. R-35, pp. 230–239.
- [6] J. Scott Provan y Michael Ball, *The complexity of counting cuts and of computing the probability that graph is connected*. SIAM, Journal on Computing, 1983, no. 12, pp. 777–788.
- [7] S. Nesmachnow, “Una versión paralela del algoritmo evolutivo para optimización multiobjetivo NSGA-II y su aplicación al diseño de redes de comunicaciones confiables,” Centro de Cálculo, Instituto de Computación, Facultad de Ingeniería, Universidad de la República, Tech. Rep. 0403, 2003.
- [8] D. Deeter and A. Smith, “Genetic algorithm design of networks considering all terminal reliability,” in *Reliability and Maintainability Symposium*, 1997, pp. 194–199.
- [9] B. Dengiz and F. Altıparmak, “Reliability optimization of computer communication networks using genetic algorithms,” in *System, Man and Cybernetics. IEEE International Conference*, vol. 5, 1998, pp. 4676–4681.
- [10] J. A. C. Falcón, *La telefonía de Internet*. Thomson, 2007.
- [11] C. J. Colbourn, L. D. Nel, T. Boffey, and D. Yates, “Network reliability and the probabilistic estimation of damage from fire spread,” *Annals of Operations Research*, vol. 50, no. 1, pp. 173–185, 1994.
- [12] A. Golro, A. S. Mohaymany, and M. Mesbah, “Reliability based investment prioritization in transportation networks,” in *Transportation Research Board 89th Annual Meeting Compendium of Papers DVD*, 2010.

- [13] B. Sanso and F. Soumis, "Communication and transportation network reliability using routing models," *IEEE Transactions on Reliability*, vol. 40, no. 1, pp. 29–38, 1991.
- [14] Héctor Cancela, "Técnicas de reducción recursiva de la varianza para la simulación de modelos estáticos y dinámicos de confiabilidad, actas de fisiología," 2000.
- [15] R. V. Slyke and H. Frank, "Network reliability analysis: part 1," *Networks*, vol. 1, pp. 279–290, 1971.
- [16] G. Fishman, *A Monte Carlo sampling plan for estimating reliability parameters and related function*. *Networks: An International Journal*, 1987, vol. 17, pp. 169–186.
- [17] A. Satyanarayana and A. Prabhakar, *New Topological Formula and Rapid Algorithm for Reliability Analysis of Complex Networks*. *IEEE Trans. Reliability*, 1985, vol. R-27, pp. 82–100.
- [18] F. Moskowitz and R. A. D. Center, *The Analysis of redundancy networks*. Rome Air Development Center, Air Research and Development Center, United States Air Force, 1958.
- [19] R. K. Wood, "Factoring algorithm for computing K-terminal network reliability," *IEEE Transactions on Reliability*, vol. R-35, no. 3, pp. 269–278, 1986.
- [20] Jean Walrand, Kallol Bagchi y George W. Zobrist, *Network Performance Modeling and Simulation*. Gordon and Breach Science Publishers, 1998.
- [21] C.L.Hyang, F. Tillman, and M. Lee, "System reliability evaluations techniques for complex/large systems," *IEEE Trans on Reliab*, vol. R-30, no. 5, 1981.
- [22] M. Grötschel, C. Monma, and M. Stoer, "Computational results with a cutting plane algorithm for designing communication networks with low connectivity constraints," Universität Augsburg, Allemagne, Institut für Mathematik, Tech. Rep. 188, 1989.
- [23] Gerardo Rubino y Bruno Tuffin, *Rare Event Simulation using Monte Carlo Methods*. WILEY, 2009.
- [24] G. S. Fishman, *Monte Carlo: concepts, algorithms, and applications*. Springer, 1996.
- [25] M. E. Khadiri and G. Rubino, "An efficient formulation of the standard Monte Carlo simulation of binary systems reliability," in *2nd Int. Conf. on Monte Carlo and Quasi-Monte Carlo Methods*, 1996.
- [26] H. Cancela, P. L'Ecuyer, M. Lee, G. Rubino, and B. Tuffin, "Refinements of a path-based efficient algorithm for network reliability estimation in the rare event case," in *Proceedings of Probability and Statistics in Science and Technology*, 2007.
- [27] H. Kumamoto, K. Tanaka, and K. Inoue, "Efficient evaluation of system reliability by Monte Carlo method," *IEEE Transactions on Reliability*, vol. R-26(5), 1977.

- [28] G. Fishman, *A Monte Carlo sampling plan for estimating network reliability*, 1986, vol. 34, no. 4, pp. 581–594.
- [29] H. Cancela and M. E. Urquhart, “Métodos Monte Carlo en la estimación de la confiabilidad de redes de comunicaciones. Estado del arte,” PEDECIBA Informática, Facultad de Ingeniería, Universidad de la República, Tech. Rep. INCO 94.06, 1994.
- [30] M. E. Khadiri and G. Rubino, “A Monte Carlo method based on antithetic variates for network reliability computations,” IRISA, France, Tech. Rep. PI626, 1992.
- [31] —, “Reliability evaluation of communications networks,” in *International Conference on Safety, Security and Reliability of Computers, SAFECOMP*, 1992.
- [32] H. Kumamoto, K. Tanaka, K. Inoue, and E. Henley, “Dagger-sampling Monte Carlo for system unavailability evaluation,” *IEEE Trans. Reliab*, vol. R-29, 1980.
- [33] C. K. Wong and M. C. Easton, “Sequential destruction method for Monte Carlo evaluation of system reliability,” *IEEE Transactions on Reliability*, 1980.
- [34] E. Dimson, “Risk measurement when shares are subject to infrequent trading,” *Journal of Financial Economics*, vol. 7, pp. 197–216, 1979.
- [35] Z. Schuster and K. Potter, “Using damage functions and total risk to quantify the impacts of climate change on extreme precipitation,” *American Geophysical Union*, 2010.
- [36] H. Kumamoto, K. Tanaka, and K. Inoue, “Efficient evaluation of system reliability by Monte Carlo method,” *IEEE Trans. Reliab*, no. R-29(2), 1980.
- [37] C. Jun and S. Ross, “System reliability by simulation: Random hazards versus importance sampling,” *Probability in the Engineering and Informational Sciences*, no. 6, 1992.
- [38] R. Karp and M. Luby, “A new Monte Carlo method for estimating the failure probability of an n-component system,” EECS Department, University of California, Berkeley, Tech. Rep. UCB/CSD-83-117, 1983.
- [39] Héctor Cancela and Mohamed El Khadiri, “A recursive variance reduction algorithm for estimating communication network reliability,” *IEEE Transactions on Reliability*, vol. 44, pp. 595–602, 1994.
- [40] H. Cancela and M. E. Khadiri, “Series-Parallel Reductiones in Monte Carlo Network Reliability Evaluation,” *IEEE Transactions on Reliability*, vol. 47, no. 2, pp. 159–164, 1998.
- [41] L. Petingi, “A diameter-constrained network reliability model to determine the probability that a communication network meets delay constraints,” *World Scientific and Engineering Academy and Society Transactions on Communications*, vol. 7, no. 6, 2008.
- [42] H. Cancela, J. Coll, D. Constanzo, and M. Rodríguez, “Modelos de confiabilidad diámetro-acotada para optimizar la profundidad y amplitud de búsqueda en redes P2P,” in *Congreso Lationamericano de Informática. CLEI*, 2006.

- [43] L. Petingi and J. Rodriguez, "Reliability of networks with delay constraints," in *Congressus Numerantium*, no. 152, 2001, pp. 117–123.
- [44] Héctor Cancela y Luis Petingi, "Diameter constrained network reliability: exact evaluation by factorization and bounds," in *International Conference on Industrial Logistics*, 2001, pp. 359–366.
- [45] S. Rai, M. Veeraraghavan, and K. Trivedi, "A survey of efficient reliability computation using disjoint products approach," *Networks, An International Journal*, vol. 25, no. 3, pp. 147–163, 1995.
- [46] H. Cancela, M. E. Khadiri, and L. Petingi, "An exact method to evaluate the source-to-terminal diameter constrained reliability of a communication network," in *International Conference on Industrial Logistics*, 2010.
- [47] —, "Polynomial-time topological reductions that preserve diameter constrained reliability of a communication network," PEDECIBA Informática, Instituto de Computación, Facultad de Ingeniería, Universidad de la República, Tech. Rep. 10-12, 2010.
- [48] Héctor Cancela and Mohamed El Khadiri, "The recursive variance-reduction simulation algorithm for network reliability evaluation," *IEEE Transactions on Reliability*, vol. 53, no. 2, 2003.
- [49] L. Murray, "Comparación entre Cinco Métodos de Monte Carlo para Estimar la Confiabilidad Diámetro Acotada de Redes de Comunicaciones," Facultad de Ciencias Exactas, Ingeniería y Agrimensura, Universidad Nacional de Rosario, Tech. Rep., 2007.
- [50] G. C. y Roger L. Berger, *Statistical Inference*. Duxbury advanced series, 2002.
- [51] H. Cancela and L. Petingi, "Domination invariant of a diameter constrained network reliability model," *Electronic Notes in Discrete Mathematics*, vol. 18, pp. 53–58, 2004.
- [52] —, "Reliability of communication networks with delay constraints: computational complexity and complete topologies," *International journal of Mathematics and Mathematical and Sciences*, pp. 1551–1562, 2004.
- [53] L. Petingi, "Combinatorial and computational properties of a diameter constrained network reliability model." ACC'08 Proceedings of the World Scientific Engineering Academy and Society International Conference on Applied Computing Conference World Scientific and Engineering Academy and Society, 2008.
- [54] H. Cancela and M. E. Khadiri, "Sensitivity of the RVR Monte-Carlo method to cutset selection strategy," in *RESIM:7^o International Workshop on Rare Event Simulation*, 2008.
- [55] —, "On the accuracy of the rvr estimator of k-terminal unreliability parameter," in *7th International Workshop on Rare Event Simulation*, 2008.
- [56] H. Cancela, M. E. Khadiri, and G. Rubino, "An efficient simulation method for k-network reliability problem," in *6th International Workshop on Rare Event Simulation*, 2006.

- [57] Schruben and Lee, *A Coverage Function for Interval Estimators of simulation response*. Management Science, January 1980, vol. 26, no. 1.
- [58] Frank J. Massey, Jr, “The Kolmogorov-Smirnov Test for Goodness of Fit,” *Journal of the American Statistical Association*, vol. 46, no. 253, pp. 68–78, 1951.
- [59] Robert Tarjan, “Depth-first search and linear graph algorithms,” *SIAM J. COMPUT*, vol. 1, no. 2, pp. 146–160, 1972.
- [60] R. Bhandari, *Survivable Networks: algorithms for diverse routing*. Springer, 1999.
- [61] L. R. Ford and D. R. Fulkerson, “Maximal flow through a network,” *Canadian Journal of Mathematics* 8, pp. 399–404, 1956.
- [62] E. W. Dijkstra, “A note on two problems in connexion with graphs,” *Numerische Mathematik*, vol. 1, pp. 269–271, 1959.
- [63] D. Alevras, M. Grötschel, and R. Wessäly, “A network dimensioning tool,” in *Preprint SC 96-49, Konrad-Zuse-Zentrum für Informationstechnik*, 1996.

Lista de figuras

2.1. Confiabilidad $s-t$ terminal, K - terminal y V - terminal	21
2.2. Grafo puente	23
2.3. Ejemplo: nodo objetivo aislado y grafo trivial C_4	24
2.4. Contracción de una arista: G y $G * e$	26
2.5. Eliminación de una arista: G y $G - e$	26
2.6. Reducción en paralelo	27
2.7. Aristas en serie	27
2.8. Ejemplo Algoritmo Factorización	32
2.9. Reducción del subgrafo G_2	32
2.10. Intervalos de los enlaces.	42
2.11. Partición de Ω_G basada en paths	52
2.12. Partición basada en minpaths	53
2.13. Partición de Ω_G basada en mincut	54
2.14. Partición basada en cortes	55
3.1. Ejemplo Grilla 3×3	60
3.2. Confiabilidad $s-t$ terminal diámetro acotada: caminos operativos	61
3.3. Confiabilidad $s-t$ terminal diámetro acotada: conjuntos de corte	62
3.4. Grafo puente	62
3.5. Reducción path-terminal	66
3.6. Muestreo multinomial	75
3.7. Ejemplo aplicación Algoritmo 15	80
5.1. Ejemplo: reducción afecta distancias	92
5.2. Ejemplo: reducción afecta distancias	93
5.3. Ejemplo: conjunto de corte	93
5.4. Grafo Completo, $n = 3, 4$	97
5.5. Red Grilla $n \times n$, $K = \{k_1, k_2, k_3, k_4\}$	98
5.6. Arpanet	99
5.7. Dodecahedro equiconfiable	100
5.8. Dodecahedro, para q variando en 2 valores, $\{q_h, q_r\}$	101
5.9. Red de transporte de ANTEL	102
5.10. Red de transporte ANTEL: región oeste y este respectivamente	102
5.11. Casos elegidos para evaluar la función de cobertura	106

5.12. Función de cobertura para grafo Puente, $Q_{\{s,t\},(n-1)}(G)$ y $q \in \{10^{-1}, 10^{-6}\}$.	107
5.13. Función de cobertura para grafo Dodecahedro, $Q_{\{s,t\},(n-1)}(G)$ y $q \in \{10^{-1}, 10^{-6}\}$	107
5.14. Función de cobertura: para red de ANTEL, región oeste, $Q_{\{s,t\},(n-1)}(G)$ y $q \in \{10^{-1}, 10^{-6}\}$	108
5.15. Grafo triángulo	109
5.16. Evolución de ER para el grafo de la Figura 5.15. Calculado en base a RVR versión multinomial.	113
7.1. Algoritmo de búsqueda en profundidad	132
9.1. Complejidad: relación entre clases.	150

Parte V
ANEXOS

Capítulo 7

Algoritmos auxiliares

Entre los algoritmos centrales para el cálculo de la medida de confiabilidad diámetro acotada se consideran operaciones en forma reiterada. Entre éstas, evaluar si una red está operativa, identificar el conjunto de corte de una red y encontrar el camino más corto entre un par de nodos. .

El algoritmo de búsqueda en profundidad [59] se aplica para evaluar el estado de una red, en el contexto de este trabajo se evalúa si los nodos objetivo están conectados.

El algoritmo de búsqueda en anchura [60] también se aplica para evaluar el estado de una red para el caso de confiabilidad diámetro acotado. El Algoritmo Ford-Fulkerson [61] permite identificar conjuntos de corte mínimo de un grafo, de forma que la eliminación de los elementos del corte hacen que la red quede en estado de falla.

El algoritmo de Dijkstra [62] se aplica para identificar el camino más corto entre dos nodos y una adaptación sencilla permite identificar el camino más corto con una cota superior que restringe el largo de los caminos buscados. El algoritmo de muestreo Multinomial se aplica en el método *RVR* para determinar el número de estados que se sortean de cada componente de una partición del espacios de estados del grafo.

7.1. Algoritmo de búsqueda en profundidad

El algoritmo de búsqueda en profundidad [59], DFS por sus siglas en inglés Depth First Search, permite recorrer todos los nodos que integran un grafo conexo generando caminos en forma recurrente. De esta manera permite distinguir si un grafo es conexo o no.

Dado un grafo $G(V, E)$, la idea es recorrer los nodos iniciando en un nodo cualquiera de G y construir caminos agregando en cada paso un nuevo nodo hasta que todos los nodos alcanzables hayan sido visitados.

La búsqueda se inicia en $s \in V$ y cada nodo tiene dos estados posibles: *Visitado* o *NoVisitado*. Si s tiene vecinos sin visitar, se elige uno de ellos, x_1 que desde este momento es *Visitado* . Desde x_1 , se identifican sus vecinos no visitados. Si existe al menos uno se elige uno de ellos a visitar, x_2 . En cambio si x_1 no tiene nodos vecinos por visitar, se vuelve a s y se buscan los vecinos sin visitar de s .

Se construye así un camino simple a un nodo de $V, \{s, \dots, x_1, x_2, \dots, t\}$. El último nodo del camino es una hoja del árbol recubridor en construcción y no tiene vecinos sin visitar. Si aún quedan nodos de G sin visitar se retrocede al vértice anterior en el camino, x , y se construye otro camino simple que nace en x e incorpora nodos aún no visitados. Si desde x no es posible construir ningún camino con nuevos nodos se sigue retrocediendo, en este caso hacia w . La aplicación del algoritmo continúa hasta construir un árbol recubridor de G . El tiempo de búsqueda empleado es lineal en V y E .

La Figura 7.1 muestra un ejemplo donde la búsqueda comienza en el nodo a y el orden de visita coincide con el orden del abecedario. El número indicado entre paréntesis refiere al orden en que ocurren las visitas. En el contexto de esta tesis, el algoritmo *DFS* se aplica para determinar si una red falla u opera. Si todos los nodos objetivos son visitados y además están a distancia D o menos dos a dos entonces la red opera.

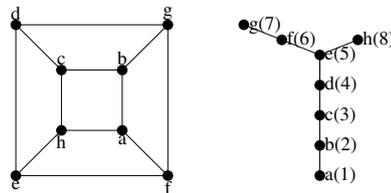


Figura 7.1: Algoritmo de búsqueda en profundidad

En el Algoritmo 16 el método parte de una red G con un conjunto de nodos objetivo K y cota de diámetro D y retorna 1 si la red opera y 0 si no. Su aplicación es sencilla y sólo requiere la especificación de un criterio que defina el orden en que se eligen los vecinos a visitar. Una opción es elegir el vecino con la menor etiqueta.

En la Línea 1 se define la variable que retorna el algoritmo iniciando en 0. En caso de que la red parámetro de entrada opere su valor se cambia a 1.

En la Línea 2 se define el vector T que contendrá los nodos visitados en cada paso, en el orden que fueron visitados: $T(1)$ indica el primer nodo visitado, $T(2)$ el segundo, etc. El vector se inicializa con el nodo u a partir del cual comienza la búsqueda y una vez finalizado el algoritmo T indicará el árbol recubridor de $K \subseteq V$. Por definición el primer elemento de *padre* y T coinciden.

En la Línea 4 se define el conjunto de nodos *NoVisitados* y en la Línea 5 se define el conjunto de nodos *Objetivo*. El primero, como indica su nombre, es el conjunto de nodos de G que no fueron visitados, y el segundo conjunto indica aquellos nodos de los que se requiere exista conectividad. En la Línea 5 se define un vector *padre* que almacena nodos predecesores de nodos visitados, de forma que el elemento i del vector *padre* indica el nodo desde que cual se visitó al nodo de la posición i de T .

En la Línea 6 se define v , que indica el nodo visitado en el último paso de la búsqueda del cual deberán identificarse vecinos por visitar. En efecto en la Línea 7 se define el conjunto *Vecinos* que contiene los nodos adyacentes a v que no pertenecen a T .

ALGORITMO 16 Búsqueda en profundidad:DFS**ENTRADA:**

grafo G ,
 nodos objetivo K ,
 confiabilidad de los enlaces \vec{r} .

SALIDA:

Variable de tipo Booleano: *Opera*. Es verdadero si el grafo está operativo.

```

1:  $Opera = 0$ 
2:  $hijo = u$ 
3:  $padre = u$ 
4:  $NoVisitados = V \setminus u$ 
5:  $Objetivo = K \setminus u$ 
6:  $v = padre$ 
7:  $Vecinos = \{Adyacencia(G, v) \setminus T\}$ 
8: Mientras No hay Vecinos y Hay Nodos por visitar
9:   Si Hay Vecinos entonces
10:     Mientras Hay Vecinos
11:        $padre = [padre, v]$ 
12:        $v = minimo(Vecinos)$ 
13:        $hijo = [hijo, v]$ 
14:        $NoVisitados = NoVisitados \setminus v$ 
15:        $Objetivo = Objetivo \setminus v$ 
16:        $Vecinos = \{Adyacencia(G, v) \setminus hijo\}$ 
17:     Fin Mientras
18:   Si no
19:      $v = padre(v)$ 
20:      $Vecinos = \{Adyacencia(G, v) \setminus hijo\}$ 
21:   Fin Si
22:   Si  $v == Raíz(1)$  entonces
23:      $HayVecinos = FALSO$ 
24:   Fin Si
25: Fin Mientras
26: Si Visité a todos los nodos Objetivo entonces
27:    $Opera = Verdadero$ 
28: Fin Si

```

Se denomina lista de adyacencia del vértice v a aquella que contiene todos los vértices w tal que $(v, w) \in E$. El conjunto de listas de adyacencia (una para cada vértice de G) se denomina estructura de adyacencia de G . En el bloque comprendido entre las Líneas 8 y 25 se indican las iteraciones, que continúan mientras queden nodos sin visitar. Iniciando por los vecinos de v (Línea 6) mientras encuentra vecino para construir el camino, se actualizan el vector $padre$ y T . En la Línea 11 se agrega al conjunto de nodos $padre$ el vector v y desde v se busca un

vecino, definido como aquél con etiqueta mínima. En la Línea 14 se actualiza el conjunto de nodos *NoVisitados* y en la Línea 15 se actualiza el conjunto *Objetivo*.

La iteración se detiene cuando no se encuentran vecinos sin visitar, lo que equivale a que la búsqueda llegó al final de un camino, a una hoja del árbol recubridor de G . En este caso se comienza a retroceder el camino generado, buscando un nodo *padre* de los ya visitados que tenga vecinos sin visitar, como indican las Líneas 19 y 20. La búsqueda continúa retrocediendo hasta que encuentra algún vecino sin visitar. Si en la búsqueda hacia atrás se continúa hasta que el nodo *Padre* coincide con su hijo en el árbol entonces se llegó al inicio de la construcción de la búsqueda en profundidad y la iteración debe terminar. Se indica este caso en las Líneas 22, 23 y 24.

Si se visitaron a todos los nodos *Objetivo* el grafo es operativo y se actualiza el parámetro de salida *Opera*.

El tiempo de ejecución de BFS es $O(|V| + |E|)$.

7.2. Algoritmo de Búsqueda en Anchura

El Algoritmo de Búsqueda en Anchura, en inglés BFS-Breadth First Search, permite recorrer los nodos que integran un grafo, y al igual que el *DFS* es un algoritmo de búsqueda “ciega” (sin información). Comenzando desde un nodo raíz el objetivo es visitar todos los nodos incorporando en forma consecutiva los vecinos de un nodo actual. Cada nodo del grafo G puede ser “visitado” o “no visitado”. Si se inicia de un nodo raíz o fuente éste se considera visitado y luego se identifican los nodos vecinos del nodo raíz que aún no fueron visitados. Como éstos vecinos son alcanzables desde la raíz se cambia su estado a “visitado” y se procede a buscar vecinos de éstos nodos que aún no fueron visitados. El proceso se repite hasta que todos los nodos fueron alcanzados o bien no hayan nodos alcanzables desde los nodos visitados.

En cada paso además calcula la distancia desde el nodo raíz a los nodos visitados. Por la forma de proceder la búsqueda sólo se visitan los nodos a distancia L si fueron visitados los nodos a distancia $L - 1$ desde el nodo raíz. Además cada nodo es visitado una única vez.

El tiempo de ejecución del Algoritmo *BFS* es $O(|V| + |E|)$.

En la Línea 1 se inicializa un vector de n componentes *Estado*, donde $|V| = n$, tal que el elemento i indica el estado del nodo i . Cada nodo tienen dos estados posibles: *Visitado* y *NoVisitado*.

En la Línea 2 se inicializa un vector de n componentes *Padre* donde el elemento i indica el nodo predecesor o padre del nodo i en el árbol de búsqueda.

En la Línea 3 se inicializa un vector de n componentes, *Distancia*, tal que el elemento i indica la distancia entre el nodo i y la raíz.

En la Línea 4 se inicializa un conjunto A donde se almacenan todos los nodos visitados. Al comenzar el algoritmo el único visitado es el nodo raíz. En las Líneas posteriores se resume un procedimiento iterativo que se repite mientras queden nodos de G por visitar.

En la Línea 6 se elige el primer nodo de A y se almacena en u . En las iteraciones éste nodo representa el nodo padre a partir del cual se identifican los nodos vecinos no visitados.

Para cada nodo vecino de u no visitado, se cambia el estado a *Visitado* y se incrementa la distancia de cada nodo vecino de u en 1. El vector *Padre* se actualiza, indicando que el nodo

u es el padre de cada uno de sus vecinos. Por último los nodos alcanzados se incorporan al vector A . Éstos pasos se resumen en las Líneas 9 a 16.

ALGORITMO 17 Búsqueda en anchura: BFS

ENTRADA:

grafo G ,
 nodos objetivo K ,
 confiabilidad de los enlaces \vec{r} ,
 nodo raíz.

SALIDA:

Variable de tipo Booleano: $Opera$. Es verdadero si el grafo está operativo.

```

1:  $Opera = 0$ 
2:  $Estado = \emptyset$ 
3:  $Padre = \emptyset$ 
4:  $Distancia = \emptyset$ 
5:  $A = s$ 
6: Mientras  $A \neq \emptyset$ 
7:    $u = A(1)$ 
8:    $A = A \setminus u$ 
9:    $V = vecinos(u)$ 
10:  Para  $i = 1$  hasta  $|V|$ 
11:    Si  $V_i$  no fue visitado entonces
12:       $Estado(V(i)) = Visitado$ 
13:       $Distancia(V(i)) = Distancia(u) + 1$ 
14:       $Padre(V(i)) = u$ 
15:       $A = A \cup V(i)$ 
16:    Fin Si
17:  Fin Para
18: Fin Mientras
19: Si  $Estado(j) == Visitado, \forall j \in K$  entonces
20:    $Opera = 1$ 
21: Fin Si

```

7.3. Algoritmo de Búsqueda en Anchura para identificar si un grafo es D -conexo

Un grafo es D -conexo si todos los nodos objetivo están comunicados por caminos de largo menor o igual a D . La adaptación del Algoritmo 17 no sólo evalúa que los nodos sean alcanzables entre sí sino que la distancia desde el nodo origen al resto de los nodos no supere la cota. Además mientras el Algoritmo 17 parte de un nodo origen desde el cual visita al resto de los nodos, la adaptación que aquí se propone considera tantos nodos origen como nodos objetivo existen. De esta forma se asegura que la distancia de todos los nodos objetivo cumple

con la restricción de distancia.

Partiendo de un nodo objetivo (nodo origen) es posible verificar si son alcanzables todos los demás nodos por caminos de largo menor o igual a la cota D . Si a partir del nodo objetivo se llega a la distancia máxima sin haber visitado todos los nodos objetivo entonces el grafo no está operativo. En el caso s - t -terminal basta aplicar esta búsqueda una vez. Si se alcanzó la cota partiendo desde s y no se visitó t el grafo está en estado de falla.

En el caso K -terminal el proceso debe repetirse para asegurar que todos los nodos de K están entre sí a una distancia no mayor a D . Si $K = \{k_1, k_2, \dots, k_{|K|}\}$ entonces debe verificarse, si k_1 es el nodo origen, que los nodos consecutivos en el conjunto sean alcanzables por caminos de largo D o menos. Luego, si esto se cumple se verifica si tomando k_2 como origen, k_3, k_4, \dots etc son alcanzables por caminos de largo D o menor. Nótese que en cada iteración el número de comparaciones se reduce en una unidad (por simetría de distancia) y, si el grafo es operativo, la última comparación que se realiza es k_{n-1} respecto de k_n . Si al menos un nodo del conjunto objetivo no es alcanzable desde algún nodo de K entonces el grafo no está operativo. El Algoritmo 18 resume estas ideas, a partir del grafo $G = (V, E)$ conjunto de nodos objetivo K y cota de diámetro D retorna verdadero si el grafo es D -conexo y falso si no. En la Línea 1 se define una variable booleana que indica si el grafo parámetro de entrada está operativo.

En la Línea 2 se define el parámetro k como el número de nodos origen de los que debe verificarse la distancia al resto de los nodos.

En la Línea 3 se inicializa Entre las Líneas que restan del algoritmo se evalúa si el grafo está operativo, de forma que se repite hasta que se identifica que el nodo no está operativo o bien hasta que se verifica que todos los nodos son alcanzables entre sí por caminos de largo menor o igual a D .

En la Línea 3 se define el parámetro *origen*, que corresponde al índice del conjunto de nodos objetivo. En cada iteración el nodo origen a partir del cual se hace una búsqueda en anchura es $K(\text{origen})$

En la Línea 5 se inicializa un vector de k componentes *Estado*, donde $|V| = n$, tal que el elemento i indica el estado del nodo i . Cada nodo tiene dos estados posibles, “visitado” o “no visitado”.

En la Línea 6 se inicializa un vector de k componentes *Padre* donde el elemento i indica el nodo predecesor o padre del nodo i en el árbol de búsqueda.

En la Línea 7 se inicializa un vector de k componentes, *Distancia*, tal que el elemento i indica la distancia entre el nodo i y la raíz.

Los vectores *Estado*, *Padre* y *Distancia* se inicializan cada vez que se modifica el nodo origen desde el que se inicia la búsqueda.

En la Línea 8 se define el estado del nodo origen como visitado y en las Líneas posteriores, de 9 a 28, se aplica la búsqueda en profundidad detallada en el Algoritmo 17 con la diferencia que sólo busca hasta que la distancia es D .

En la Línea 9 se inicializa el vector Q de nodos visitados, inicializado con el nodo origen.

ALGORITMO 18 Búsqueda en anchura en redes con distancia acotada entre nodos**ENTRADA:**

grafo G ,
 nodos objetivo K ,
 confiabilidad de los enlaces \vec{r} ,
 cota D .

SALIDA:

Variable de tipo Booleano: *Opera*. Es verdadero si el grafo está operativo.

```

1: Opera = VERDADERO
2:  $k = |K|$ 
3: origen = 1
4: Mientras opera y origen  $\leq k$ 
5:   Estado =  $\emptyset$ 
6:   Padre =  $\emptyset$ 
7:   Distancia =  $\emptyset$ 
8:    $K(\textit{origen}) = \textit{Visitado}$ .
9:    $Q = K(\textit{origen})$ 
10:  nodos = 1
11:   $B = \emptyset$ 
12:  Para  $j = 1$  hasta  $D$ 
13:    Mientras  $Q \neq \emptyset$ 
14:       $u = Q(1)$ 
15:       $Q = Q \setminus u$ 
16:       $V = \textit{vecinos}(u)$ 
17:      Para  $i = (\textit{origen} + 1)$  hasta  $k$ 
18:        Si  $K(i)$  no fue visitado entonces
19:          Estado  $K(i)$  = visitado
20:           $\textit{Distancia}(K(i)) = \textit{Distancia}(u) + 1$ 
21:           $\textit{Padre}(K(i)) = u$ 
22:           $B = B \cup K(i)$ 
23:          Si  $K(i)$  es un nodo objetivo aún no visitado entonces
24:            nodos = nodos + 1
25:          Fin Si
26:        Fin Si
27:      Fin Para
28:    Fin Mientras
29:     $Q = Q \cup B$ 
30:  Fin Para
31:  Si No visité todos los nodos objetivo entonces
32:    Opera = FALSO
33:  Fin Si
34:  origen = origen + 1
35: Fin Mientras

```

En la Línea 10 se define un contador *nodos* que refiere al número de nodos visitados. En la Línea 11 se define un vector *B* que se inicia vacío y en las iteraciones acumula los vecinos visitados.

En la Línea 12 se indican *D* iteraciones con el objetivo de identificar los nodos que están a distancia *D* o menos del nodo origen. El vector *Q* contiene los nodos visitados, y en particular en la primera iteración sólo consta del nodo origen. En los pasos siguientes se actualiza con los vecinos visitados.

En la *j*-ésima iteración se evalúa cuáles son los nodos alcanzables a distancia *j* del origen, actualizando los vectores de *Distancia* y *Estado* de los nodos y el *Padre* o predecesor en el árbol de búsqueda. La primera iteración evalúa todos los nodos adyacentes al nodo origen y que por lo tanto están a 1 paso de distancia. En la siguiente iteración se visitan los nodos adyacentes de los ya visitados, identificando así los nodos a dos pasos del origen. Cada vez que se visita un nodo objetivo aún no visitado el contador *nodos* se incrementa en una unidad. Cuando finalizan las *D* repeticiones, en las Líneas 29 a 31 se verifica si todos los nodos objetivos fueron alcanzados a *D* pasos o menos, dado un nodo origen. Si la condición $nodos < |K|$ la red falla y la búsqueda termina. En caso contrario se repite el proceso partiendo desde otro nodo objetivo.

7.4. Algoritmo de Ford-Fulkerson

El algoritmo de Ford-Fulkerson tiene aplicaciones diversas. Una de ellas identificar el flujo máximo a través de una red con capacidades asociadas a los enlaces. Se define un flujo a enviar a través de los enlaces de la red con valores asociados de capacidad, de forma que el máximo flujo que puede enviarse por un enlace es la capacidad que éste tiene. En el caso particular que la capacidad de todos los enlaces es 1, la aplicación del algoritmo permite identificar el conjunto de corte mínimo de una red.

De la aplicación del algoritmo de Ford Fulkerson se obtiene el flujo máximo que puede enviarse a través de una red, valor que coincide con el corte de capacidad mínima de la red.

Teorema 7.4.1 *Min cut- max flow* [61]

El flujo máximo entre un nodo fuente y uno terminal de un grafo dirigido es igual al corte de menor capacidad

Previo a detallar el algoritmo de Ford- Fulkerson se formaliza la definición de “flujo de una red” y “camino aumentable”.

Definición 7.4.2 Sea $G = (V, E)$ una red con vértices V y enlaces E , capacidad de los enlaces $c : E \rightarrow \mathbb{R}^+$, y $\{s, t\} \in V$, nodos fuente y terminal respectivamente. Diremos que $f : E \rightarrow \mathbb{R}^+ \cup \{0\}$ es un **flujo** en G si:

$$0 \leq f(e) \leq c(e), \forall e \in E \quad (7.1)$$

$$\sum_{e \rightsquigarrow v} f(a) = \sum_{v \rightsquigarrow e} f(a), \forall v \in V \setminus \{s, t\} \quad (7.2)$$

La función de flujo tiene una restricción de capacidad y otra de conservación de flujo, de forma que no se envía a través de la red más que su capacidad y todo el flujo que llega a un nodo es la misma cantidad que sale de éste (exceptuando los nodos fuente y terminal).

Definición 7.4.3 Una secuencia de vértices-arcos $v_0, e_1, v_1, e_2, \dots, e_n, v_n$ es un **camino aumentable** si

- $e_i = v_{i-1}v_i$ o $e_i = v_i v_{i-1}, \forall i = 1, 2, \dots, n$.
- $e_i \neq e_j \forall i \neq j$.
- $e_i = v_{i-1}v_i$ y $f(e_i) < c(e_i)$ o $f(e_i) > 0$.

Se denomina *gap* de e_i a $g(e_i) = c(e_i) - f(e_i)$ si $e_i = v_{i-1}v_i$ o $f(e_i)$ si no.

Definición 7.4.4 Si P es un camino aumentable $P = \{s = v_0, e_1, \dots, e_n, v_n = t\}$ y $g = \min_{\text{gap}}(e_i)$, se denomina **flujo aumentado** de f según P a la función $f' : E \rightarrow \mathbb{R}^+$: $f' = \begin{cases} f(e) & \text{si } e \notin P \\ f'(e) = f(e) + g & \text{si } e = v_i v_{i+1} \\ f'(e) = f(e) - g & \text{si } e = v_{i+1} v_i. \end{cases}$

El costo de los enlaces se define en función de la confiabilidad elemental de forma que el corte de costo mínimo coincida con el corte de mayor probabilidad de falla. El costo del i -ésimo enlace en E es $-\log(q_i)$ y se cumple

$$\max \prod_{i=1}^{|C|} q_i = \max \ln \left(\prod_{i=1}^{|C|} q_i \right) = \max \sum_{i=1}^{|C|} \ln(q_i) = \min \sum -\ln(q_i).$$

La etiqueta $(-, +\infty)$ indica que a través del nodo fuente ingresa tanto flujo como sea necesario. La etiqueta $(fuente^+, \Delta(v))$ indica que el flujo desde el nodo *fuente* hacia v puede aumentarse en $\Delta(v)$ unidades. La etiqueta $(u^+, \Delta(w))$ indica que el flujo existente en el nodo w puede aumentarse en $\Delta(w)$ tomada del nodo u . La etiqueta $(u^-, \Delta(w))$ indica que al disminuir el flujo de w a u , el total del flujo que sale de w a los vértices etiquetados puede reducirse en $\Delta(w)$. Ésta cantidad puede usarse para incrementar el flujo desde w hasta vértices no etiquetados.

El procedimiento de etiquetado está diseñado para identificar el flujo máximo que puede enviarse a través de una red. Dado que existen nodos con múltiples nodos adyacentes el resultado no es único pudiendo existir distintas combinaciones que permitan maximizar el flujo. Cuando se realiza el procedimiento de etiquetado y al culminar el nodo terminal no tiene etiqueta entonces el proceso de búsqueda termina, se alcanzó el flujo máximo.

En esta instancia existen nodos no etiquetados y nodos etiquetados, de forma que los enlaces que conectan los nodos del primer conjunto con los del segundo tienen flujo igual a la capacidad del enlace.

ALGORITMO 19 Etiquetado**ENTRADA:**

grafo G ,

nodos objetivo: fuente y terminal,

matriz de capacidad de los enlaces: CAP .

SALIDA:

lista de etiquetas asignadas a los nodos: $Etiqueta$.

1: $flujo = 0 \forall e \in E$

2: $Etiqueta(fuente) = (-, +\infty)$

3: $vecino = Vecinos(Gdata, fuente)$

4: **Para** $i = 1$ **hasta** $|vecino|$

5: $v = vecino(i)$

6: **Si** $CAP(fuente, v) > flujo(fuente, v)$ **entonces**

7: $\Delta(v) = CAP(fuente, v) - flujo(fuente, v)$

8: $Etiqueta(v) = (fuente^+, \Delta(v))$

9: **Fin Si**

10: **Fin Para**

11: **Mientras** Existe $e = (u, w)$ tal que $u, u \neq fuente$ tiene etiqueta, w no tiene etiqueta

12: **Si** $CAP(u, w) > flujo(u, w)$ **entonces**

13: $\Delta(w) = \min\{\Delta(u), CAP(u, w) - flujo(u, w)\}$

14: $Etiquetas(w) = (u^+, \Delta(w))$

15: **Fin Si**

16: **Fin Mientras**

17: **Mientras** Existe $e = (w, U)$ tal que $u, u \neq fuente$ tiene etiqueta, w no tiene etiqueta

18: **Si** $flujo(w, u) > 0$ **entonces**

19: $\Delta(w) = \min\{\Delta(u), CAP(u, w) - flujo(u, w)\}$

20: $Etiquetas(w) = (u^-, \Delta(w))$

21: **Fin Si**

22: **Fin Mientras**

ALGORITMO 20 MinCut-MaxFlow**ENTRADA:**

grafo G ,
 nodos objetivo: fuente y terminal,
 matriz de capacidad de los enlaces: CAP .

SALIDA:

Mincut de G : C .

- 1: $Etiquetas = Etiquetado(G, fuente, terminal, CAP)$
- 2: **Si** Nodo terminal tiene etiqueta **entonces**
- 3: **Si** $Etiquetas(terminal) = (x^+, \Delta(terminal))$ **entonces**
- 4: $flujo(x, terminal) = flujo(x, terminal) + \Delta(terminal)$
- 5: **Fin Si**
- 6: $x = NodoEnEtiqueta(terminal)$
- 7: $v = NodoEnEtiqueta(x)$
- 8: **Mientras** $x \neq fuente$
- 9: **Si** $Etiqueta(x) = (v^+, \Delta(x))$ ($\Delta(x) \geq \Delta(terminal)$) **entonces**
- 10: $flujo(v, x) = flujo(v, x) + \Delta(terminal)$
- 11: **Fin Si**
- 12: **Si** $Etiqueta(x) = (v^-, \Delta(x))$ ($\Delta(x) \geq \Delta(terminal)$) **entonces**
- 13: $flujo(x, v) = flujo(x, v) - \Delta(terminal)$
- 14: $x = NodoEnEtiqueta(v)$
- 15: $v = NodoEnEtiqueta(x)$
- 16: **Fin Si**
- 17: **Fin Mientras**
- 18: **Fin Si**
- 19: **Si** Nodo terminal no está etiquetado **entonces**
- 20: $Etiquetas = ConjNodosEtiquetados(G)$
- 21: $NoEtiquetas = V \setminus ConjNodosEtiquetados(G)$
- 22: El conjunto de corte C es el conjunto de enlaces que separan $Etiquetas$ de $NoEtiquetas$.
- 23: $FlujoMaximo = \sum_{e_i \in C} CAP(e_i)$
- 24: **Fin Si**

7.5. Camino más corto con restricciones de largo: Algoritmo Dijkstra adaptado

El algoritmo de Dijkstra [62] permite identificar el camino más corto entre dos nodos de un grafo y el algoritmo que identifica el camino más corto con restricciones de largo es una adaptación de éste.

Dado un grafo $G = (V, E)$ y matriz de costos $Costo$ asociado a los enlaces, donde $Costo_{i,j}$ indica el costo asociado al enlace que une los nodos $i \neq j \in V$ es $1 \forall i, j$, el problema es resoluble en tiempo polinomial [63].

Sea L la cota superior para el largo de los caminos entre dos nodos de G . Se define $d_i(v)$ como la distancia más corta entre un nodo origen s y v usando a lo sumo i enlaces. Se define además

el conjunto U de nodos no etiquetados, que son los nodos a los que no se ha llegado por un camino de largo menor o igual a L desde s . Se define también $R \subseteq U$ como el conjunto de nodos no etiquetados que pueden alcanzarse usando nodos etiquetados intermedios: $R = \{v \in U : \exists k \leq L, d_k(v) < \infty\}$. La etiqueta asignada corresponde a la distancia más corta desde el nodo s a v .

El Algoritmo 21 resume el procedimiento para identificar el camino más corto entre un par de nodos de un grafo G . A partir del grafo, los nodos que se desean conectar y la cota superior para el largo del camino el algoritmo retorna el camino más corto entre dichos nodos. En la Línea 1 se inicializa el conjunto de nodos no etiquetados U , que contiene a todos los nodos de G excepto el nodo origen s . En la Línea 2 se inicializa el conjunto R que contiene a todos los nodos de G vecinos de s . En la Línea 3 se inicializa un vector *predecesor* que permite reconstruir el camino más corto empezando desde t hacia s .

Para cada v , en las Líneas 4 a 11 se define la distancia de s a v . Si s y v coinciden entonces $d_i(v) = 0$. Si la arista (s, v) existe entonces la distancia es 1 y si no hay enlace que los conecte entonces la distancia es ∞ .

Definidas las etiquetas iniciales de los nodos se considera el conjunto de nodos R , que indica los nodos alcanzables desde s dado los nodos ya visitados (etiquetados).

El bloque comprendido entre las Líneas 15 y 32 se repite mientras hay nodos alcanzables no etiquetados.

En la Línea 16 se define v , el siguiente nodo a etiquetar, que es el nodo más cercano a s que no tiene etiqueta.

Identificado v se actualizan las etiquetas. En cada nodo de U se evalúa qué camino es más largo, el camino de s a U_k o el camino de s a v y de v a U_k . Si la segunda opción permite comunicar los nodos en menos pasos entonces se actualiza la etiqueta. Éstos pasos son los indicados en las Líneas 21 y 28, donde además se almacena el nodo predecesor a U_k .

En las Líneas 29 a 31, si hay nodos etiquetados con distancia menor que ∞ entonces son nodos alcanzables y se incorporan al conjunto R .

Una vez que se etiquetaron todos los nodos de V resta identificar la distancia mínima entre s y t . Si la etiqueta más pequeña vale ∞ entonces no existe camino en G que conecte ambos nodos. En cambio si existe un valor positivo y menor a ∞ entonces se reconstruye el camino en función de la información almacenada en *predecesor*.

ALGORITMO 21 Camino más corto largo acotado**ENTRADA:**grafo G ,cota superior de largo de los caminos: L ,nodo origen: s .**SALIDA:**camino más corto entre s y cualquier nodos de $V - s$, de largo menor a L .

```

1:  $U = V - s$ 
2:  $R = \{v \in V / (s, v) \in E\}$ 
3:  $predecesor = s$ 
4: Para  $v \in V$ 
5:   Si  $e = (s, v) \in E$  entonces
6:      $d_i(v) = 1$ 
7:   Si no
8:     Si  $s = v$  entonces
9:        $d_i(v) = 0$ 
10:    Si no
11:       $d_i(v) = \infty$ 
12:    Fin Si
13:  Fin Si
14: Fin Para
15: Mientras  $R \neq \emptyset$ 
16:    $v = \text{NodoMasCercano}(s)$ 
17:    $U = U - \{v\}$ 
18:    $R = R - \{v\}$ 
19:   Para  $k = 1$  hasta  $|U|$ 
20:     Para  $j = (i + 1)$  hasta  $L$ 
21:       Si  $(d(v, j - 1) + \text{Dist}(v, U_k)) < d(U_k, j)$  entonces
22:          $predecesor(U_k) = v$ 
23:       Fin Si
24:        $d(U_k, j) = \min(d(U_k, j), d(v, j - 1) + \text{Dist}(v, U_k))$ 
25:     Fin Para
26:     Si  $p(U_k, L) < \infty$  entonces
27:        $R = R \cup U_k$ 
28:     Fin Si
29:   Fin Para
30: Fin Mientras
31:  $\text{DistanciaMinima} = \min(\text{dist}(s, t))$ 
32: Si  $\text{DistanciaMinima} = \infty$  entonces
33:    $\text{camino} = \emptyset$ 
34: Si no
35:    $\text{camino} = \text{ReconstruyoCamino}(s, t, \text{predecesor})$ 
36: Fin Si

```

7.6. Muestra aleatoria Multinomial

En el Algoritmo 15 propuesto en el Capítulo 3 para la estimación de la medida $R_{K,D}(G)$ se incluye el sorteo de una muestra Multinomial. El Algoritmo 23 considera algunas de las propiedades de la distribución multinomial [24].

La distribución es de la forma:

$$f(z_1, z_2, \dots, z_r) = N! \prod_{i=1}^r \left(\frac{p_i^{z_i}}{z_i!} \right)$$

donde (Z_1, Z_2, \dots, Z_r) tienen distribución Multinomial y se cumple

- $0 < p_i < 1$
- $\sum_{i=1}^r p_i = 1$
- $z_i \geq 0, z_i \in \mathbb{N}$
- $\sum_{i=1}^r z_i = N$

La distribución marginal de Z_j es binomial:

$$f(z_j) = C_{z_j}^N p_j^{z_j} (1 - p_j)^{n - z_j},$$

y la distribución de Z_j condicionada a $(Z_1, Z_2, \dots, Z_{j-1}) = (z_1, z_2, \dots, z_{j-1})$

$$f(z_j | (Z_1, Z_2, \dots, Z_{j-1}) = (z_1, z_2, \dots, z_{j-1})) = C_{z_j}^{N - \sum_{i=1}^{j-1} z_i} w_j^{z_j} (1 - p_j)^{n - \sum_{i=1}^{j-1} z_i},$$

con $w_j = \frac{p_j}{1 - \sum_{i=1}^{j-1} p_i}, j = 2, 3, \dots, r$.

Antes de detallar el algoritmo para generar una realización de la distribución Multinomial se presenta el Algoritmo 22 que especifica como generar la distribución Binomial.

Dado los parámetros N y $p \in \mathbb{R}$ el algoritmo retorna Z , el número de éxitos en los N sorteos.

En la Línea 1 se inicializa Z en 0 y en la Línea 2 se inicia un contador i y se itera hasta que $i = N$. Entre las Línea 4 se sortea una variable Uniforme(0,1) U . Si $U < p$ la variable Z se incrementa una unidad.

El Algoritmo 23 resume la obtención de una muestra multinomial. Dado el tamaño de sorteos N y el vector de probabilidades $p \in \mathbb{R}^r$, en las primeras líneas inicializa y define los parámetros r, M, j y p . El valor r corresponde al número de componentes de p y del vector multinomial. El valor M se actualiza durante el procedimiento e indica el número de elementos que resta sortear de N . El parámetro j indica el valor de p y Z que se considera en cada iteración. Por último q se actualiza en cada paso y representa el denominador del parámetro $w_j = \frac{p_j}{q}$.

El vector Z se inicializa con todos los componentes en 0 y es el parámetro que retorna el Algoritmo 23. En las Líneas 6 a 11 se indica el proceso iterativo. En cada paso se sortea una

muestra Binomial(M, w_j) mientras quedan elementos para sortear. La búsqueda termina cuando $M = 0$, se sortearon N elementos, o cuando $j = r + 1$, se generaron las r entradas del vector Z . En el primer caso el vector Z tendrá componentes finales igual a 0. Cada entrada Z_j se sortea de acuerdo con una distribución condicionada a las $j - 1$ valores de Z_i ya sorteados según el Algoritmo 22.

ALGORITMO 22 Muestra Binomial

ENTRADA:

tamaño de muestra N ,
 probabilidad de éxito p .

SALIDA:

muestra binomial de tamaño N : Z .

- 1: $Z = 0$
 - 2: $i = 1$
 - 3: **Mientras** $i \leq N$
 - 4: $U = \text{Uniforme}(0, 1)$
 - 5: **Si** $U < p$ **entonces**
 - 6: $Z = Z + 1$
 - 7: **Fin Si**
 - 8: $i = i + 1$
 - 9: **Fin Mientras**
-

ALGORITMO 23 Muestra Multinomial

ENTRADA:

tamaño de muestra N ,
 vector de probabilidad de éxito p .

SALIDA:

muestra multinomial: Z .

- 1: $M = N$
 - 2: $q = 1$
 - 3: $r = \text{length}(p)$
 - 4: $j = 1$
 - 5: $Z = \text{ceros}(r, 1)$
 - 6: **Mientras** $M * (r - j + 1) > 0$
 - 7: $Z_j = \text{Binomial}(M, \frac{p_j}{q})$
 - 8: $M = M - Z_j$
 - 9: $q = q - p_j$
 - 10: $j = j + 1$
 - 11: **Fin Mientras**
-

Capítulo 8

Trabajo en el cluster

Cada trabajo implementado en el cluster se envía utilizando PuTTY. Es el script utilizado tiene formato .sh, que indica que es un archivo ejecutable en linux.

Del total de líneas que contiene el script, las únicas líneas que se modifican para cada corrida implementada en esta tesis son:

- La Línea 2 que indica el nombre del trabajo.
- La Línea 5 que indica el tiempo máximo de cómputo que se solicita, en horas (h), minutos (m) y segundos (s).
- La Línea 51 que indica el nombre del archivo .m que se debe ejecutar en MATLAB.

```
1 #!/bin/bash
2 # Nombre del trabajo
3 #PBS -N TRABAJO
4 # Requerimientos
5 #PBS -l nodes=1,walltime=h:m:s
6 # Cola
7 #PBS -q publica
8 # Working dir
9 #PBS -d /home/bertinat
10 # Output path
11 #PBS -e /home/bertinat
12 #PBS -o /home/bertinat
13 # Correo electronico
14 #PBS -M XXXX@fing.edu.uy
15 # Email
16 #PBS -m abe
17 # n: no mail will be sent.
18 # a: mail is sent when the job is aborted by the batch system.
19 # b: mail is sent when the job begins execution.
20 # e: mail is sent when the job terminates.
```

```
21 #PBS -V
22 echo Job Name: $ PBS _ JOBNAME
23 echo Working directory: $ PBS _ O_ WORKDIR
24 echo Queue: $ PBS _ QUEUE
25 echo Cantidad de tasks: $PBS _ TASKNUM
26 echo Home: $ PBS_ O _ HOME
27 echo Puerto del MOM: $ PBS _ MOMPORT
28 echo Nombre del usuario: $ PBS_ O _ LOGNAME
29 echo Idioma: $ PBS _O _ LANG
30 echo Cookie: $ PBS _ JOBCOOKIE
31 echo Offset de numero de nodos: $ PBS _ NODENUM
32 echo Shell: $ PBS_ O _ SHELL
33 #echo JobID: $ PBS _O _ JOBID
34 echo Host: $ PBS _O _ HOST
35 echo Cola de ejecucion: $ PBS _ QUEUE
36 echo Archivo de nodos: $ PBS _ NODEFILE
37 echo Path: $ PBS _ O _ PATH
38 echo
39 cd $PBS_O_WORKDIR
40 echo Current path:
41 pwd
42 echo
43 echo Nodos:
44 cat $PBS_NODEFILE
45 echo
46 # Define number of processors
47 echo Cantidad de nodos:
48 NPROCS= $\prime$ wc -l < $ PBS _ NODEFILE $\prime$
49 echo $ NPROCS
50 echo
51 matlab -r NOMBRETRABAJO
```

Capítulo 9

Glosario

1. **#-P**: Conjunto de los problemas de conteo asociados a los problemas de decisión en el conjunto NP . Un problema pertenece a la clase #-P si existe una máquina de Turing no determinista que lo resuelve en tiempo polinomial.
2. **Árbol recubridor**: Si $G = (V, E)$ es un grafo no dirigido, un subconjunto K de V es un recubrimiento de G si para cada arista (a, b) de G , a o b está en K . El conjunto K es un recubrimiento mínimo de G si $K - x$ no cubre a G , $\forall x \in K$.
3. **Árbol de Steiner**: El problema del árbol de Steiner es un caso particular del Problema General de Steiner (GSP), cuando $r_{i,j} = 1$, $\forall i, j \in T$. Dado un grafo $G = (V, E)$, un subconjunto de nodos terminales $T \subseteq V$, una matriz de costos reales no negativos $C = \{c_{i,j}\}_{(i,j) \in E}$ y una matriz de enteros positivos $R = \{r_{i,j}\}_{i,j \in T}$, el objetivo del GSP es encontrar un subgrafo $H \subseteq G$ de costo mínimo tal que $\forall i, j \in T$ existan al menos $r_{i,j}$ caminos nodo-disjuntos comunicando i con j en H .
4. **Arborescencia**: Un grafo $G = (V, E)$ dirigido es una arborescencia con raíz s si todo punto $v \in V$ distinto de s tiene grado de entrada igual a uno, s tiene grado de entrada 0 y no hay circuitos.
5. **Clase P (Polinomial)**: Conjunto de problemas de decisión resolubles en tiempo polinomial respecto del tamaño de la instancia por una máquina secuencial determinista.
6. **Clase NP (No determinista polinomial)**: Conjunto de problemas de decisión donde puede verificarse la correctitud de una solución en tiempo polinomial por una máquina secuencial determinista.
 - a) **NP-completo**: Un problema es NP-completo si es NP y NP-difícil.
 - b) **NP-difícil**: Un problema de decisión L es NP-difícil si $\forall L' \in NP$ es reducible polinomialmente a L .
7. **$D - K$ -árboles**: Árbol recubridor de $K \subseteq V$ cuya distancia entre todo par de nodos no supera una cota D .

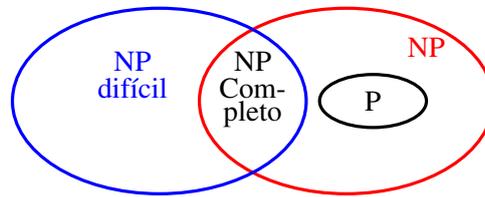


Figura 9.1: Complejidad: relación entre clases.

8. **Diámetro:** En un grafo, la distancia entre dos vértices es el menor número de aristas de un recorrido entre ellos. El diámetro es la mayor distancia entre dos puntos del mismo.
9. **Dominancia con signo:** Se obtienen como la resta entre el número de formaciones pares y el número de formaciones impares.
10. **Estimador puntual:**
 - **Consistencia:** Un estimador es consistente si converge casi seguramente al valor real con probabilidad 1 cuando el tamaño de la muestra crece a ∞ (converge en probabilidad). Formalmente, si $W_n = W_n(X_1, \dots, X_n)$ es una sucesión de estimadores, W_n es una sucesión consistente de estimadores del parámetro θ si para cada $\epsilon > 0$ y cada $\theta \in \Theta$, $\lim_{n \rightarrow \infty} \text{Prob}_\theta(|W_n - \theta| < \epsilon) = 1$.
 - **Eficiencia:** Mientras que la consistencia se asocia a la convergencia del estimador, la eficiencia refiere a la varianza asintótica del estimador. Una sucesión de estimadores es asintóticamente eficiente si para un parámetro $\tau(\theta)$ si $\sqrt{n}[W_n - \tau(\theta)] \rightarrow n(0, v(\theta))$ en distribución, donde $v(\theta) = \frac{[(\tau'(\theta))^2]}{E_\theta((\frac{\partial}{\partial \theta} \log(f(x|\theta)))^2)}$.
11. **Formación F de G:** Conjunto de $D - K$ -árboles cuya unión es el conjunto de aristas de G.
12. **Formación par(impar):** Se dice que una formación es par (impar) si está formada por un número par (impar) de árboles.
13. **Grafo Conexo:** Un grafo es conexo si cualquier par de nodos de G está conectado mediante un camino de G .
14. **Intervalo de Confianza:** Un intervalo de confianza de un parámetro θ es cualquier par de funciones $L(X), U(X)$ de una muestra $X = (x_1, x_2, \dots, x_n)$ tal que $L(X) \leq U(X) \forall x \in X$. El intervalo aleatorio $[L(X), U(X)]$ es un estimador por intervalo [50]. La probabilidad $\text{Pr}(\theta \in [L(X), U(X)]|\theta)$ es la probabilidad de que el verdadero valor del parámetro pertenezca al intervalo y se denomina **probabilidad de cobertura** [50]. El **coeficiente de confianza** del intervalo se define como el ínfimo de la probabilidad de cobertura: $\inf_\theta \text{Pr}(\theta \in [L(X), U(X)]|\theta)$ [50].

15. **Sesgo:** El sesgo de un estimador puntual W de un parámetro θ es la diferencia entre el valor esperado de W y θ , $Sesgo_{\theta}(W) = E_{\theta}(W) - \theta$.
Un estimador cuyo sesgo es 0 se dice que es insesgado y cumple $E_{\theta}(W) = \theta, \forall \theta$ [50].