



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN

MÉTRICAS DE ROBUSTEZ PARA EL ANÁLISIS DE REDES

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN CIENCIAS, MENCIÓN
COMPUTACIÓN

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL EN COMPUTACIÓN

FERNANDO GABRIEL MORALES CORREA

PROFESOR GUÍA:
JAVIER BUSTOS JIMÉNEZ

MIEMBROS DE LA COMISIÓN:
MAÍRA MARQUÉS SAMARY
JOSÉ PIQUER GARDNER
MARCIA PAIVA MOREIRA

Este trabajo ha sido parcialmente financiado por NIC CHILE RESEARCH LABS

SANTIAGO DE CHILE
2017

MÉTRICAS DE ROBUSTEZ PARA EL ANÁLISIS DE REDES

La dependencia de la gente, las empresas y el gobierno con los sistemas de telecomunicaciones se ha vuelto muy fuerte y vital en la mayoría de los casos. Una falla en estos sistemas, como por ejemplo Internet, puede ocasionar pérdidas en muchos sentidos paralizando casi cualquier tipo de industria. Así, estudiar estas fallas y fenómenos ha atraído a muchos investigadores a proponer modelos para detectarlas, estudiarlas y desarrollar acciones de mitigación y prevención. Uno de estos enfoques, es la medición de robustez de estos sistemas usando teoría de grafos, modelándolo abstractamente como elementos interconectados entre sí.

El objetivo de este trabajo se enfoca en la evaluación y mejora de este sistema bajo la mirada de las redes complejas. Específicamente, sobre la robustez de la red física del Internet chileno, que dada la geografía del país es muy vulnerable tanto a desastres de factores naturales o humanos.

Dado que el campo de las redes complejas es extremadamente extenso en la comunidad científica, es necesario estudiar los diversos sistemas y modelos propuestos, de tal manera a elegir aquel que más se adecúe a la problemática mencionada anteriormente. En detalle, se necesita una métrica adecuada para medir la robustez del sistema. Este trabajo consiste en un estudio exhaustivo de métricas de robustez para luego aplicarlas al contexto chileno.

La metodología de este trabajo consiste en aplicar una técnica muy usada en el ámbito científico llamada estudio de mapeo sistemático. Esta revisión, bajo algunos antecedentes, responde la siguiente pregunta de investigación, **qué métricas existen para estudiar la robustez en redes complejas**, realizando una búsqueda en diversos repositorios académicos.

Los resultados de este trabajo indican que la mayoría de la comunidad científica se ha concentrado en problemas de mayor escala, que no están sujetas a la particular geografía del territorio chileno. Otra razón por la cual aquellas métricas no son útiles para nuestro caso es que la mayoría darían resultados poco favorables dado que la red es intrínsecamente **poco** robusta.

Las mejoras propuestas indican correlaciones favorables sobre una métrica en particular, sobre una topología similar a menor escala, el **impacto nodal de Wiener** (no presente dentro de los resultados del mapeo sistemático), que aparte de poseer una buena conectividad como requisito, toma en cuenta tanto las distancias (latencias) de la red, como la calidad de sus caminos alternativos (en caso de falla).

Finalmente, como trabajo futuro principal se aplicarán estas mejoras a la red real, tomando en cuenta las distancias geográficas y sus costes asociados, así como la evaluación correspondiente por la métrica elegida.

A mi familia, polola y amigos

Tabla de Contenido

| | |
|--|-----------|
| Índice de Tablas | ix |
| Índice de Ilustraciones | x |
| 1. Introducción | 1 |
| 1.1. Definición del Problema | 2 |
| 1.1.1. Contexto | 2 |
| 1.1.2. Problema | 3 |
| 1.1.3. Pregunta de investigación | 3 |
| 1.2. Objetivos del trabajo | 3 |
| 1.2.1. Objetivo general | 3 |
| 1.2.2. Objetivos específicos | 4 |
| 1.3. Metodología de trabajo | 4 |
| 2. Estudio de Mapeo Sistemático | 5 |
| 2.1. Introducción | 5 |
| 2.1.1. Ventajas y Desventajas | 5 |
| 2.1.2. Características de los Estudios de Revisión Sistemática | 6 |
| 2.2. Proceso de Revisión | 6 |
| 2.3. Protocolo de Estudio de Mapeo Sistemático | 7 |
| 2.3.1. Antecedentes | 7 |
| 2.3.2. Preguntas de estudio | 7 |
| 2.3.3. Estrategia | 8 |
| 2.3.4. Criterios de selección de estudios | 9 |
| 2.3.5. Proceso de selección de estudios | 10 |
| 2.3.6. Extracción de datos | 10 |
| 2.3.7. Análisis de datos | 10 |
| 2.3.8. Método de publicación de los resultados | 11 |
| 2.4. Resumen de los resultados | 11 |
| 2.4.1. Conceptos asociados a métricas | 11 |
| 2.4.2. Aplicaciones | 11 |
| 2.4.3. Coste computacional | 12 |
| 2.4.4. Estadísticas | 13 |
| 3. Métricas encontradas | 18 |
| 3.1. Introducción | 18 |
| 3.1.1. Teoría de grafos | 19 |

| | |
|---|-----------|
| 3.1.2. Redes Complejas | 21 |
| 3.2. Métricas existentes | 22 |
| 3.2.1. Métricas basadas en eliminación | 22 |
| 3.2.2. Métricas basadas en caminos | 23 |
| 3.3. Aplicaciones | 25 |
| 3.3.1. Transporte | 25 |
| 3.3.2. Red Eléctrica | 27 |
| 3.3.3. Telecomunicaciones | 28 |
| 3.3.4. Gestión de cadenas de suministro | 30 |
| 3.3.5. Biología y medicina | 31 |
| 3.3.6. Redes sociales | 32 |
| 3.3.7. Redes complejas generales | 33 |
| 4. Caso de Estudio | 37 |
| 4.1. Backbone red Chilena | 37 |
| 4.1.1. Métricas elegidas | 39 |
| 4.1.2. Primera mejora de la red | 39 |
| 4.1.3. Segunda mejora de la red | 41 |
| 4.1.4. Discusión | 42 |
| 5. Conclusiones y trabajo futuro | 47 |
| 5.1. Conclusiones | 47 |
| 5.2. Trabajo futuro | 48 |
| 6. Bibliografía | 49 |
| 7. Anexos | 57 |
| 7.1. Consultas realizadas | 57 |
| 7.2. Tabla comparativa de categorías | 60 |

Índice de Tablas

| | |
|--|----|
| 2.1. Criterios de población, intervención, comparación, resultado y contexto. | 7 |
| 2.2. Cantidad de artículos por motor de búsqueda. | 9 |
| 2.3. Ejemplo de información de las respuestas de la pregunta 1. | 10 |
| 2.4. Ejemplo de información de las respuestas de la pregunta 2. | 10 |
| 2.5. Lista de las métricas encontradas por el mapeo sistemático. | 17 |
| 4.1. Comparación de valores de las métricas para las soluciones adicionando 3 enlaces. | 41 |
| 4.2. Valores de métricas de la solución escogida. | 43 |
| 7.1. Tabla comparativa de categorías por artículo. | 61 |
| 7.2. Tabla comparativa de categorías por artículo. | 62 |

Índice de Ilustraciones

| | | |
|------|--|----|
| 2.1. | Diagrama de búsqueda. | 8 |
| 2.2. | Frecuencia de las métricas encontradas dividida por fecha. | 13 |
| 2.3. | Frecuencia de las métricas encontradas dividida por categoría. | 13 |
| 2.4. | Frecuencia de las métricas encontradas dividida por fecha y categoría. | 14 |
| 2.5. | Diagrama que indica la frecuencia de los costes computacionales por métrica. | 14 |
| 2.6. | Diagrama que indica la frecuencia de los temas aplicados por métrica. | 15 |
| 2.7. | Diagrama que indica la frecuencia de los editoriales por métrica. | 15 |
| 4.1. | Red Universitaria Nacional (REUNA), es una corporación integrada por universidades, centros de investigación y grupos astronómicos internacionales. Los elementos en color rojo, llamados <i>bridges/cut-points</i> corresponden a aquellos elementos cuya eliminación desconecta la red. | 38 |
| 4.2. | Soluciones agregando 3 enlaces. | 40 |
| 4.3. | El rendimiento de cada solución según cada métrica seleccionada a minimizar con VNS. (MIN POSSIBLE) se refiere a los mínimos globales de cada métrica por separado (encontradas por “fuerza bruta”, no es aplicable a grafos más grande). Notar que una solución que tenga el mínimo global para una métrica no necesariamente minimiza globalmente otras. | 43 |
| 4.4. | Valores de las métricas por cada solución, ordenadas por el impacto nodal de Wiener. | 44 |
| 4.5. | Valores de las métricas por cada solución, ordenadas por el impacto nodal de Wiener. En escala logarítmica. | 45 |
| 4.6. | Solución encontrada minimizando MAX_NWI , con 4 enlaces adicionales. | 46 |

Capítulo 1

Introducción

Chile es un país altamente sísmico con una geografía diversa, cuya larga y estrecha forma no hace natural que se tenga una infraestructura de telecomunicaciones redundante. La enorme mayoría de las fibras ópticas que interconectan el país pasan a metros de la Ruta 5, y la probabilidad de cortes masivos es demasiado alta. En varias emergencias ya sean naturales o artificiales, el país ha presenciado como el servicio de Internet y telefónico se ve interrumpido sin aviso en grandes superficies del territorio, dejando aislados a miles de compatriotas.

La telefonía se ha estructurado de forma que resulta imposible que opere bien en caso de congestión, las cuales siempre ocurren en consecuencia de las emergencias. En cambio Internet es una red resiliente por diseño, preparada para corregir fallas parciales y para continuar en operación mientras existan caminos alternativos disponibles. Por otro lado, la ciudadanía, el gobierno y las empresas se han ido volviendo cada vez más dependientes de la infraestructura de Internet, y la enorme penetración de las redes sociales la hace una infraestructura crítica para casos de emergencia. La red Internet debiera estar disponible y operativa el 100% del tiempo, incluso durante un gran desastre, de modo que la gente logre comunicarse, tranquilizarse y organizarse sin dar espacio para el pánico que genera el no tener noticias de nuestros seres queridos.

Las empresas de telecomunicaciones invierten en sus redes privadas de datos, para tener infraestructura que les permita vender conectividad en todo el país. Las grandes ciudades cuentan con acceso vía fibra óptica mientras que el mundo rural solo cuenta con enlaces inalámbricos, que son mucho más limitados.

Esta vulnerabilidad y desigualdad de acceso en la red es cada vez más acentuada a medida que nos alejamos más de la ciudad de Santiago. Hasta el punto en que es imposible hablar de robustez de redes en los pueblos de Chile.

1.1. Definición del Problema

Existen varias maneras (y conceptos) para medir vulnerabilidades en redes. Bajo el contexto de la *teoría de grafos*, Las *redes complejas*, refieren a una red (grafo) que posee ciertas propiedades estadísticas y topológicas no triviales que no ocurren en redes simples, por ejemplo, distribuciones de grado que siguen leyes de potencia, estructuras jerárquicas, estructuras comunitarias.

Un ejemplo de estas redes pueden ser las redes sociales formadas por Facebook¹ o Twitter², cuyos nodos son interpretados como entes personales o corporativas y las conexiones pueden ser interpretadas como una relación de “amigos” (relación simétrica), o “siguiendo” (relación no necesariamente simétrica).

Otro ejemplo, es una red de cableado eléctrico o bien de fibra óptica. Esta última, se utiliza extensivamente para modelar matemáticamente la topología de la Internet chilena.

La *robustez*, inclusive sólo en palabras de las redes complejas, es definida de muchas formas. La definición más usada corresponde a la habilidad de resistir fallas y perturbaciones. Otro concepto existente es la *resiliencia*, que corresponde a la habilidad de proveer y mantener un nivel aceptable de servicio frente a fallas ante una operación normal. Cabe mencionar que estas definiciones no son lo suficientemente objetivas para medirlas matemáticamente, por lo que, dado cada contexto y tipo de red (y muchos factores más), existen muchas formas de medir la robustez.

1.1.1. Contexto

La Subsecretaría de Telecomunicaciones (SUBTEL³), organismo dependiente del Ministerio de Transportes y Telecomunicaciones y encargado de coordinar, promover, fomentar y desarrollar las telecomunicaciones en Chile, ha hecho grandes esfuerzos financiando proyectos de telecomunicaciones en todo el país, sin embargo la necesidad de la población no solo es tener conectividad en tiempos normales, sino que esta sea resiliente, que resista con cierto grado de certeza las inclemencias propias de nuestra naturaleza.

Dado esta situación, es importante estudiar la robustez de nuestra red. En el año 2015 CORFO⁴ aprobó un proyecto⁵ que planea estudiar y evaluar la resiliencia de la infraestructura de la Internet chilena. Dentro de los alcances de este proyecto, se pretende desarrollar una métrica basada en teoría de grafos para medir matemáticamente qué tan resiliente es la Internet chilena y, en base a esos resultados, proponer mejoras para que el país entero este más preparado para emergencias.

¹<https://www.facebook.com/>

²<https://twitter.com/>

³<http://www.subtel.gob.cl/>

⁴<http://www.corfo.cl/>

⁵Proyecto CORFO 15BPE-47225: “Estudio y recomendaciones sobre la resiliencia de la infraestructura del internet chileno”

1.1.2. Problema

Como anteriormente mencionado, bajo el contexto de las redes complejas, se han desarrollado diversas métricas que miden de alguna forma en particular la robustez de una red [20, 24, 26, 35, 48]. Pero todavía no existe una visión global completa y clara de éstos indicadores. Esto último ocurre porque una métrica se puede desarrollar en distintas disciplinas, como la biología o la física [4, 30, 32].

Existen estudios que resumen parcialmente estas métricas, pero no necesariamente están actualizados ni comprenden diversas definiciones (o nociones) de robustez. Un pequeño ejemplo es mencionado por Sydney *et al.* [74], quienes nombran 2 acercamientos de robustez en redes complejas. La *clásica*, se basa en la topología y fundamentos matemáticos de teoría de grafos (clique, grado, camino más corto); y la *contemporánea* se refiere a la habilidad de una red para mantener su flujo total (o degradarse suavemente) ante la eliminación de nodos y aristas, tomando en cuenta los posibles servicios de la red.

1.1.3. Pregunta de investigación

Este trabajo de tesis busca responder, mediante un estudio extensivo en la literatura científica, la siguiente pregunta: **qué métricas existen para estudiar la robustez en redes complejas**. A partir de este estudio, se realizará un análisis de la red chilena dada sus características (pocos proveedores de Internet y redes concentradas) presentando una recomendación de las métricas aplicables.

De esta manera, el país estará mejor preparado para seguir conectado ante emergencias y desastres de toda índole, además de contar con infraestructura de alta capacidad en sectores que necesiten redundancia en ruteo de datos nacional.

1.2. Objetivos del trabajo

Los objetivos planteados para el desarrollo de este trabajo de tesis se presentan a continuación. Se incluye tanto el objetivo general como los objetivos específicos.

1.2.1. Objetivo general

El objetivo general consiste en desarrollar un mapeo sistemático acerca de las métricas existentes en la literatura científica. Luego a partir de este estudio, realizar un análisis del Internet chileno con el objetivo de presentar una recomendación de las métricas aplicables.

1.2.2. Objetivos específicos

- Realizar un listado de las métricas encontradas.
- Clasificar según origen y aplicación las métricas encontradas.
- Encontrar grupos de métricas similares en base a sus aplicaciones y que tipo de conceptos de teoría de grafos son utilizados.
- Estudiar su aplicación a la Internet chilena presentando una recomendación de las métricas útiles para cada caso.

1.3. Metodología de trabajo

El desarrollo de la presente tesis siguió la siguiente metodología de trabajo. Ésta se divide en etapas, las cuales a su vez se subdividen en tareas.

Primera etapa: desarrollo del protocolo de mapeo sistemático

1. Definir los pasos del estudio de mapeo sistemático.
2. Preparar un breve informe acerca de los pasos del estudio.

Segunda etapa: realizar estudio de mapeo sistemático

3. Realizar las consultas a los repositorios académicos.
4. Aplicar criterios de selección.
5. Extraer y analizar los datos de cada uno de los artículos aprobados.
6. Etiquetar en coste, categoría y aplicación las métricas extraídas.

Tercera etapa: análisis de la red chilena

7. Seleccionar y aplicar métricas sobre la topología de la red física chilena.
8. Proponer mejoras en base a estos resultados.

Cuarta etapa: conclusiones

9. Concluir respecto a los resultados del estudio de mapeo sistemático y la robustez de la red chilena.

Capítulo 2

Estudio de Mapeo Sistemático

2.1. Introducción

Dado que nuestra pregunta de investigación requiere realizar un estudio imparcial y extensivo sobre un tema amplio en la literatura científica, la comunidad investigadora ha desarrollado formas de estudios sistemáticos. El que se adecúa a nuestra situación corresponde al *estudio de mapeo sistemático*, descrito en [40] como una forma de estudio que utiliza una metodología previamente definida para identificar, analizar e interpretar todas las evidencias disponibles sobre alguna pregunta de investigación de tal forma que sea imparcial y (hasta cierto grado) repetible.

La mayoría de las investigaciones comienzan con un estudio (de alguna forma) previo a la literatura científica. Sin embargo, si este estudio no es exhaustivo ni objetivo, posee poco valor científico. Éste es el motivo principal por el cual se realizan estudios sistemáticos, ya que resumen el trabajo reciente (o no tan reciente) de una manera objetiva. Por ejemplo, el estudio debe realizarse siguiendo una estrategia predefinida. En particular, un estudio sistemático debe identificar y reportar (a toda costa) todos los estudios concernientes.

2.1.1. Ventajas y Desventajas

Las ventajas de un estudio de revisión sistemática incluyen:

- Una metodología bien definida que hace menos probable que los resultados no sean imparciales.
- Pueden dar información acerca de los efectos de algún fenómeno a través de un amplio rango de escenarios y metodologías empíricas. Si los estudios encuentran resultados consistentes, la revisión sistemática provee evidencia de que el fenómeno es robusto y transferible. Si los estudios proveen resultados inconsistentes, se pueden estudiar las variaciones.
- En el caso de los estudios cuantitativos, es posible combinar datos usando técnicas

meta-analíticas. Aumentando la probabilidad de encontrar efectos que estudios menores puedan no detectar.

La mayor desventaja es que requieren un esfuerzo considerablemente mayor que estudios de literatura tradicionales. Además, como existe mayor capacidad de meta-análisis, es posible detectar pequeñas parcialidades como relaciones consistentes.

2.1.2. Características de los Estudios de Revisión Sistemática

Algunas de las características que diferencian un estudio de revisión sistemática de uno convencional son:

- Requieren de un protocolo de revisión que especifica (al menos) una pregunta de investigación y los métodos a utilizar durante la revisión.
- Se basan en una estrategia de búsqueda predefinida que busca detectar la mayor cantidad de literatura concerniente a la pregunta de investigación.-
- Se documenta la estrategia de búsqueda para que los lectores puedan verificar la rigurosidad, completitud y reproducción del proceso (Considerando que las búsquedas en librerías digitales son casi imposibles de repetir).
- Requieren criterios de inclusión y exclusión específicos para verificar el potencial de cada artículo.
- Especifican la información obtenida por cada estudio primario incluyendo el criterio de calidad con el cual se evaluó el artículo.
- Es un prerrequisito de un análisis cuantitativo.

2.2. Proceso de Revisión

Un estudio de mapeo sistemático consiste en definir los siguientes pasos:

- Antecedentes: la necesidad de una revisión sistemática nace del requerimiento de los investigadores en resumir información existente. Antes de realizar la revisión, se debe evaluar si es necesaria. De esta manera, encontrar la pregunta adecuada y correspondiente al problema en cuestión.
- Preguntas de estudio: las preguntas de investigación que la revisión intenta responder. Idealmente, la estructura de la pregunta debe incluir los criterios de *población* (gente o tema general que comprende a los antecedentes descritos), *intervención* (metodología/herramientas/tecnología que estudian una problemática específica), *comparación* (metodología/herramientas/tecnología con la cual la intervención será comparada), *resultado* (resultados o mejoras deseadas y esperadas con respecto a los antecedentes descritos) y *contexto* (Temas y participantes en el cual la comparación será realizada).
- Estrategia: forma y lugar donde la búsqueda de información será realizada.

- Criterio(s) de selección de estudio: criterios de inclusión y exclusión que serán aplicados a los resultados de la estrategia de búsqueda.
- Proceso de selección de estudios: metodología por la cual se aplicarán los criterios definidos anteriormente.
- Extracción de datos: metodología y descripción de los datos a extraer del artículo.
- Análisis de datos: preguntas relacionadas que responden la pregunta principal de investigación.
- Método(s) de publicación de los resultados: artículo o forma de divulgación por la cual los resultados fueran descritos.

2.3. Protocolo de Estudio de Mapeo Sistemático

A continuación, se describirán cada una de las partes del protocolo usado en este trabajo de tesis.

2.3.1. Antecedentes

Bajo la necesidad de encontrar nuevas formas eficientes de medir robustez de la Internet, y bajo el aspecto de las redes complejas, se formula la siguiente pregunta: **¿Qué métricas existen para estudiar la robustez en redes complejas?**

2.3.2. Preguntas de estudio

En la tabla 2.1 se describen los criterios de población, intervención, comparación, resultado y contexto.

Tabla 2.1: Criterios de población, intervención, comparación, resultado y contexto.

| Criterio | Descripción |
|--------------|---|
| Población | Redes complejas. |
| Intervención | Métricas de robustez. |
| Comparación | Diferentes métricas de robustez, resiliencia y percolación que podamos encontrar. |
| Resultado | Descripción y categoría de la métrica. |
| Contexto | Industria de telecomunicaciones y transporte, políticas públicas, investigación de operaciones. |

A continuación, se presentan las preguntas asociadas a la pregunta principal.

- ¿Qué métricas estudia el paper?
- ¿Cuáles son las aplicaciones (de existir) que el paper propone o estudia?

Cada una de estas preguntas fue desglosada en varias preguntas específicas (detalladas en la descripción de análisis de datos) de tal manera que puedan ser respondidas de la manera más objetiva y clara posible.

2.3.3. Estrategia

Se realiza una búsqueda manual en los siguientes repositorios:

- ACM Digital Library.
- IEEE Xplore (IEEE Electronic Library - IEL).
- ScienceDirect.
- Web of Science.
- PLoS (Public Library of Science).
- Nature.

Se seleccionaron estos repositorios porque son los más relevantes en los tópicos de las redes complejas.

Un diagrama de la búsqueda realizada es mostrado en la figura 2.1. Las operaciones lógicas $X \text{ AND } Y$ y $X \text{ OR } Y$ son para buscar papers que contengan los términos X e Y , o X o Y respectivamente. La operación $X \text{ PRE}/n \ Y$ es para buscar el término X sucedido por Y en n palabras. Los procesos de *stemming*¹ de los motores de búsqueda son considerados. Se buscan estos términos tanto en el *abstract* como el título del paper. Las consultas exactas realizadas en cada uno de los repositorios están descritas en la sección 7 Anexos.

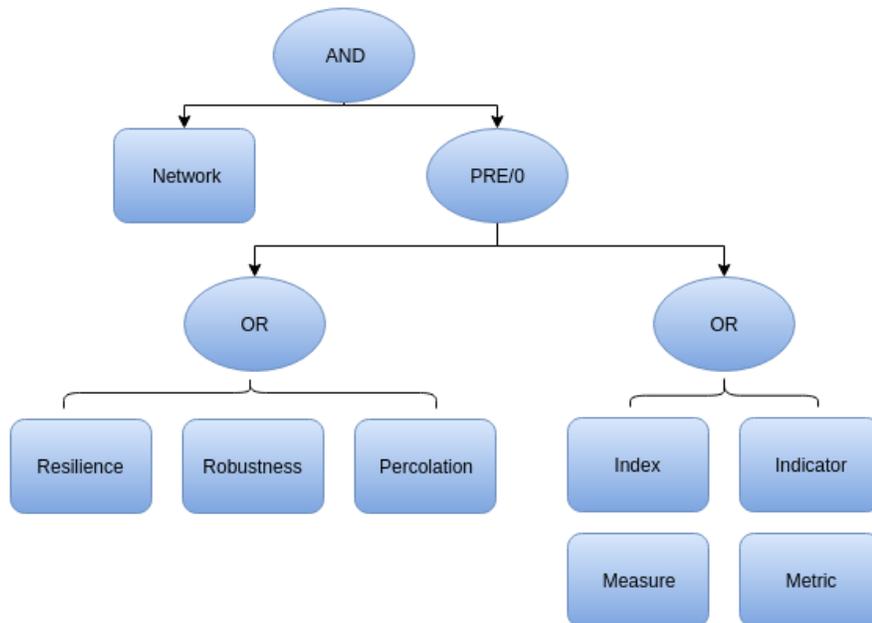


Figura 2.1: Diagrama de búsqueda.

¹Es un método para reducir una palabra a su raíz.

En la tabla 2.2, se muestran la cantidad de artículos resultados de la consulta a cada uno de los motores de búsqueda.

Tabla 2.2: Cantidad de artículos por motor de búsqueda.

| Motor de búsqueda | Cantidad de artículos |
|--------------------------|------------------------------|
| ACM Digital Library | 2 |
| IEEE Xplore | 11 |
| ScienceDirect | 6 |
| Web of Science | 192 |
| PLoS | 6 |
| Nature | 2 |

2.3.4. Criterios de selección de estudios

Dados los resultados de una búsqueda se aplican los criterios de exclusión/inclusión para la selección de los estudios a analizar.

Criterios de inclusión

El paper:

- Esta en inglés.
- Es un estudio primario (journal).
- No es un estudio *survey*.
- Fue publicado entre Enero de 2005 y Diciembre 2016.
- Está publicado en una revista científica.
- Está disponible online.
- Tiene como tópico las redes complejas.

Criterios de exclusión

El paper:

- Tiene como tópico las redes neuronales.
- No declara resultados concluyentes.
- No estudia las métricas de robustez.

2.3.5. Proceso de selección de estudios

Los resultados de la búsqueda fueron ordenados por año y fuente. Luego, analizados por dos investigadores, que leyeron solamente el título y el *abstract* aplicando los criterios de inclusión/exclusión correspondientes.

La lista de papers aprobados y rechazados se mantienen en una lista. Los rechazados estarán listados junto con el motivo de su rechazo. En el caso de no existir un acuerdo en la inclusión un paper, un tercer investigador revisó el artículo y decidió si incluirlo o no.

2.3.6. Extracción de datos

De cada paper se extraerán el título, autor, abstract, DOI y año de publicación.

2.3.7. Análisis de datos

En esta sección de la metodología propuesta, se definen como se responderán las preguntas de investigación. Junto con la información que aporta la pregunta y una posible respuesta. En la tabla 2.3 se presenta el análisis de la primera pregunta propuesta: ¿Qué métricas estudia el paper? La tabla 2.4 presenta el análisis de la segunda pregunta propuesta: ¿Cuáles son las aplicaciones (de existir) de la métrica, propuestas por el estudio?

Tabla 2.3: Ejemplo de información de las respuestas de la pregunta 1.

| Pregunta | Información | Posible respuesta |
|--|----------------|---|
| ¿Cuál es el nombre de la métrica? | Nombre | Betweenness Centrality |
| Cuál es la definición formal de esta métrica? | Definición | Cantidad de caminos más cortos que pasan por un nodo en particular |
| A qué categoría pertenece esta métrica? | Categoría | Robustez |
| Qué interpretación tiene un valor de esta métrica? | Interpretación | Generalmente, un bajo valor representa menos vulnerabilidad para aquel nodo |

Tabla 2.4: Ejemplo de información de las respuestas de la pregunta 2.

| Pregunta | Información | Posible respuesta |
|---|-------------|-------------------|
| ¿Cuáles son las aplicaciones (de existir) de la métrica, propuestas por el estudio? | Aplicación | Medicina |

2.3.8. Método de publicación de los resultados

Los resultados de este estudio son de interés del proyecto mencionado en la sección Introducción. Son publicados en el presente informe de trabajo de tesis.

2.4. Resumen de los resultados

En esta sección se describirán brevemente los resultados obtenidos por la revisión sistemática.

2.4.1. Conceptos asociados a métricas

Basados en el origen y conceptos usados por cada métrica estudiada, los siguientes grupos se detectaron. Nótese que una métrica puede pertenecer a varios grupos o ninguno dependiendo del caso. Cada grupo se presenta como $X (Y) [Z]$, donde X es el concepto asociado, Y su traducción al español y Z su etiquetación asociada para posteriores referencias. En la sección 7 Apéndice se presenta una tabla indicando las categorías de cada uno de los artículos que propusieron una métrica.

- *Connectivity* (Conectividad) [*Con*]: Se refiere a que tan conectados están los vértices y/o componentes del grafo.
- *Degree* (Grado) [*Deg*]: Se refiere al número de aristas incidentes a un vértice.
- *Percolation* (Percolación) [*P*]: Se refiere a la teoría la cual asocia a cada vértice con un valor representando la probabilidad o estado de infección del vértice. Es usado recientemente para modelar la robustez en redes.
- *Centrality* (Centralidad) [*Cen*]: Se refiere a la importancia de algún elemento del grafo con respecto a otros o bien al grafo completo.
- *Distance* (Distancia) [*D*]: Se refiere al tamaño del camino más corto entre dos vértices.
- *Spectral* (Espectral) [*Sp*]: Se refiere a la teoría la cual analiza los vectores y valores propios que representa una matriz. Por ejemplo la matriz de adyacencia del grafo.
- *Varies* (Varía) [*V*]: Se refiere a que el tema puede variar debido a la definición de la métrica.

2.4.2. Aplicaciones

Cada métrica es aplicada en el contexto de alguno de los siguientes temas.

- Infraestructuras críticas.
 - Transporte: conjunto de viajes de aerolíneas, líneas de ferrocarril o metro.

- Red eléctrica: redes interconectadas donde se transfiere energía en forma de electricidad y pueden existir de estaciones generadoras de electricidad, actuando de proveedores, junto a los consumidores, que pueden ser máquinas o centros demanda.
- Telecomunicaciones: redes que se dedican al intercambio de señales de cualquier naturaleza, típicamente electromagnéticas, que contengan signos, sonidos, imágenes o, en definitiva, cualquier tipo de información que se desee comunicar a cierta distancia. Un ejemplo muy común corresponde a la red Internet.
- Cadenas de suministros: una abstracción de una red constituidas generalmente de proveedores, distribuidores y consumidores. Un proveedor es un elemento que elabora un bien, el cual es transmitido por el distribuidor y finalmente recibido por el consumidor.
- Biología y medicina: por lo general, redes de interacción de proteínas, donde un nodo es una proteína y un enlace representa una reacción química.
- Redes sociales: relaciones humanas, ya sea formada por interacciones de trabajo, amistad o cualquier índole.
- Temas generales de redes complejas: redes que no poseen un mayor significado además de la existencia de sí misma.

2.4.3. Coste computacional

Cada métrica tiene un coste computacional (en tiempo) asociado dentro de las siguientes opciones, dependiendo de n y m :

- Polinomial (*POLY*): El coste se puede representar como una función proporcional a un polinomio $P(n, m)$.
- Exponencial (*EXP*): El coste se puede representar como una función proporcional a una exponencial $e(x, y)$.
- No determinado (*ND*): No se pudo determinar el coste.
- Depende de la métrica (*DEPENDS*): Debido a la definición de la métrica, el coste varía.

2.4.4. Estadísticas

De un total de 128 artículos leídos extrajeron 57 métricas. La figura 2.2 muestra cuales fueron los años donde más se definieron métricas (2011).

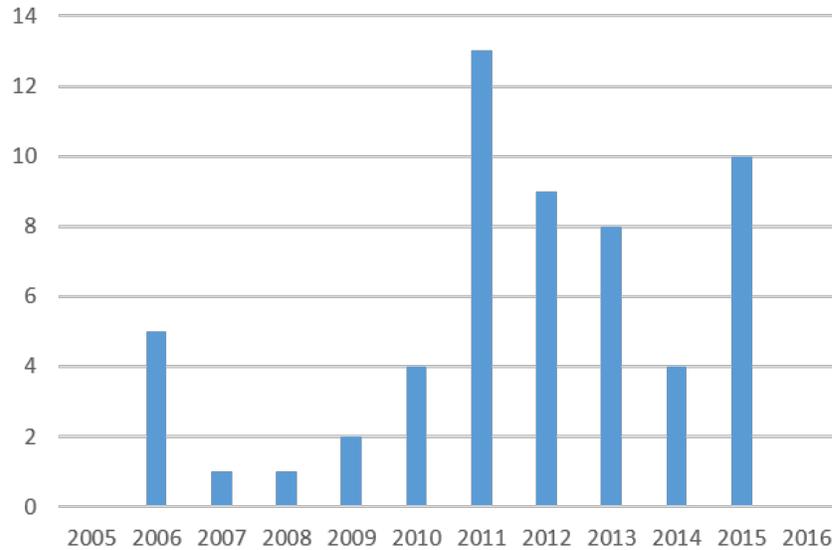


Figura 2.2: Frecuencia de las métricas encontradas dividida por fecha.

La figura 2.3 muestra que tan frecuentes fueron las categorías. Es importante notar que existen métricas que pertenecen a dos categorías y por ende la suma de las frecuencias corresponde a un número mayor a la cantidad total de métricas. Por raro que parezca no se encontraron métricas para los años 2005 ni 2016.

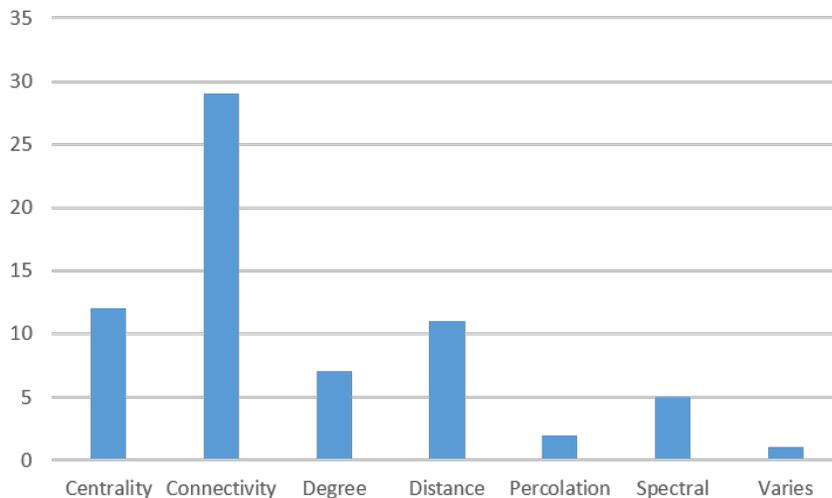


Figura 2.3: Frecuencia de las métricas encontradas dividida por categoría.

La figura 2.4 muestra la misma frecuencia sobre las fechas, pero también divididas por categoría. Se puede observar que antes del 2010 no se encontraron métricas relacionadas con la centralidad.

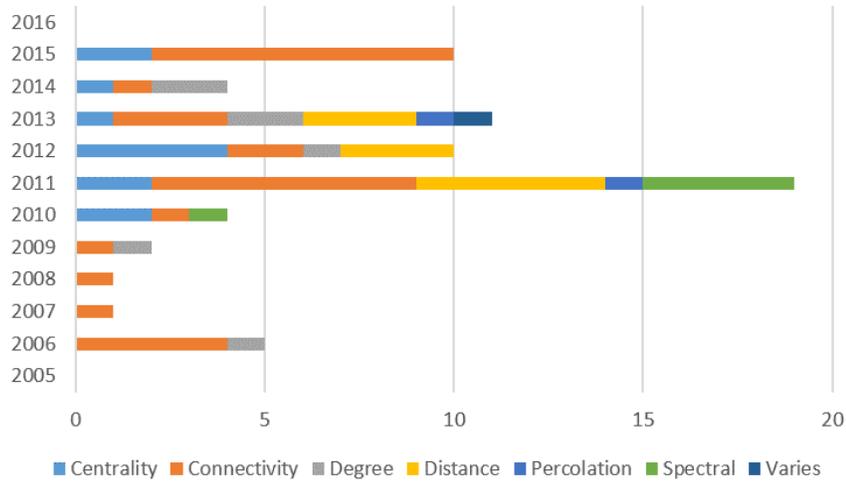


Figura 2.4: Frecuencia de las métricas encontradas dividida por fecha y categoría.

La figura 2.5 indica las proporciones de los costes con respecto al total de métricas. Cabe notar que las métricas con coste exponencial se vuelven poco prácticas a la hora de calcularla en grafos de tamaño medio.

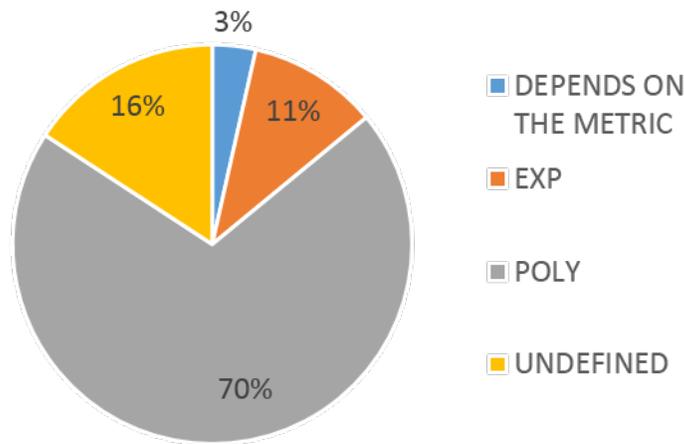


Figura 2.5: Diagrama que indica la frecuencia de los costes computacionales por métrica.

La figura 2.6, indica que la mayoría de las métricas es estudiada en temas generales de teoría de grafos, seguida por estudios de telecomunicaciones.

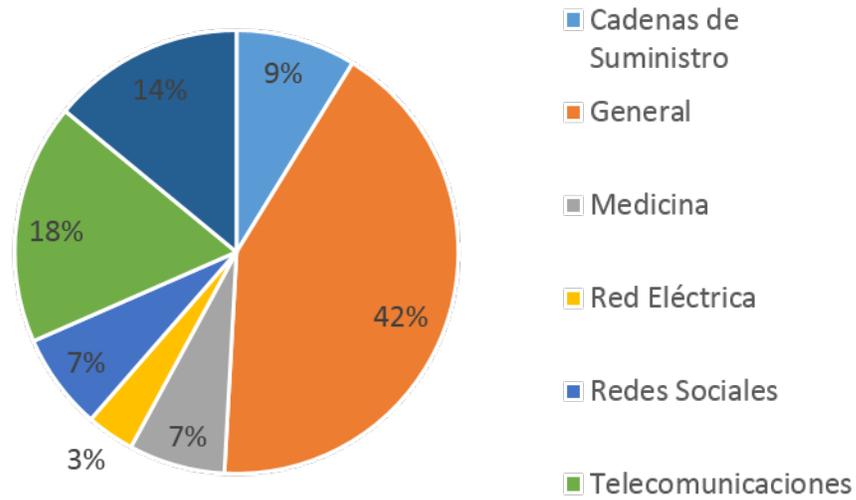


Figura 2.6: Diagrama que indica la frecuencia de los temas aplicados por métrica.

Finalmente la figura 2.7 muestra cuales fueron las principales editoriales. Dentro de la categoría otras destacan: *Springer, Taylor & Francis, National Acad Sciences, BioMed Central Ltd.*

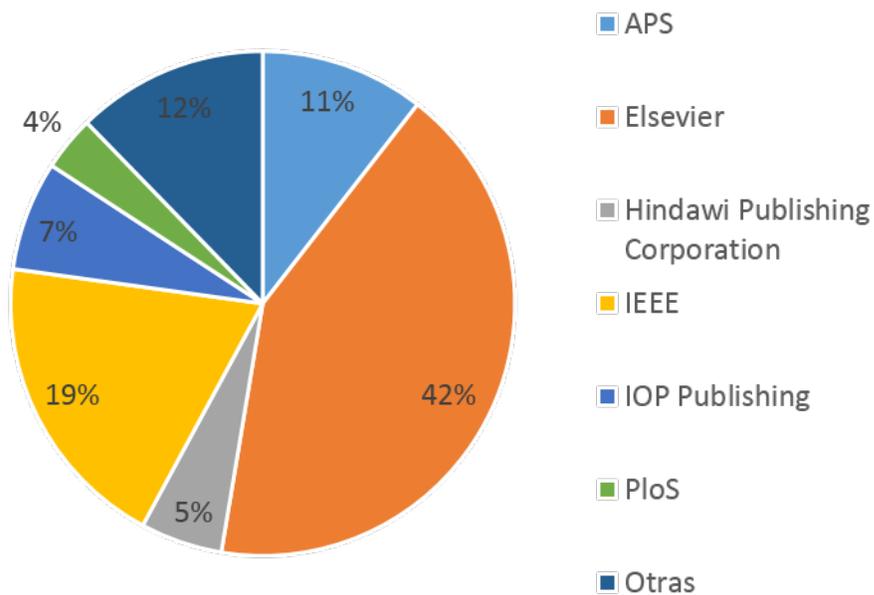


Figura 2.7: Diagrama que indica la frecuencia de los editoriales por métrica.

Como resumen, la tabla 2.5 muestra la lista de métricas encontradas con las cuales se realizaron todas las estadísticas anteriores.

| Nombre métrica | Cita |
|--|------|
| Relative betweenness | [66] |
| Betweenness Interference | [66] |
| Fragmentation | [12] |
| Reconstructability Coefficient | [46] |
| Free Energy Rank | [22] |
| Entropy Rank | [22] |
| Link Robustness Index | [90] |
| Community Robustness | [47] |
| Pairwise Disconnectivity Index | [61] |
| Resilience Factor | [65] |
| Preferential Perturbation Score | [62] |
| Entropy of the Degree Distribution | [81] |
| Perturbation Score | [62] |
| Maximum Perturbation Score | [62] |
| Average Perturbation Gradient | [62] |
| Node Load | [82] |
| Global Functionality Loss | [91] |
| Random Resilience Score | [75] |
| Viral Conductance | [89] |
| Targeted Resilience Score | [75] |
| Fragility | [45] |
| Dynamic Robustness Metric | [45] |
| Network Trip Robustness | [57] |
| Network Robustness Metric | [42] |
| Covering Degree | [64] |
| Covering Index | [64] |
| Effective Graph Resistance | [43] |
| Geographical Path Diversity | [14] |
| Effective Geographical Path Diversity | [14] |
| Normalized Aggregated Remaining Flow | [14] |
| Aggregated Remaining Flow | [14] |
| Compensated Total Graph Geographical Diversity | [14] |
| Total Graph Geographical Path Diversity | [14] |
| Robustness Measure 1 (No Name) | [80] |
| Conectivity Robustness Function | [87] |
| Efficiency Function | [87] |
| Relative Entropy | [33] |
| Node/Edge Resilience | [63] |
| Node/Edge Criticality | [78] |
| Network Criticality | [78] |
| Supply Availability | [93] |

| | |
|------------------------------|------|
| Network Connectivity | [93] |
| Best Delivery Efficiency | [93] |
| Average Delivery Efficiency | [93] |
| Natural Connectivity | [85] |
| Sensitivity | [19] |
| Deterioration | [7] |
| Temporal Efficiency | [67] |
| Bridgeness | [13] |
| Local Natural Connectivity | [88] |
| Subgraph Centrality | [86] |
| Random-Robustness Index | [76] |
| Unique Robustness Measure | [69] |
| Percolation Centrality | [59] |
| Supply Chain Index | [60] |
| Generalized Robustness Index | [50] |
| Relative Area Index | [58] |

Tabla 2.5: Lista de las métricas encontradas por el mapeo sistemático.

Capítulo 3

Métricas encontradas

En este capítulo se describirá la teoría relacionada a las redes, también llamadas grafos y las llamadas redes complejas. Teoría por la cual se modelan los problemas de redes sociales, ingenieriles, computacionales, entre otras aplicaciones. Posteriormente se describirán las métricas encontradas, aquellas que fueron propuestas como nuevas métricas en los estudios, resultado de la revisión sistemática descrita anteriormente.

3.1. Introducción

Un sistema que ejerce, ya sea, manteniendo o al menos restaurando la operabilidad bajo fallas o ataques es considerado un sistema *resiliente*. Este trabajo de tesis explora las definiciones dadas en la literatura con el objetivo de cuantificar la *resiliencia* de un sistema interconectado. Cuando la resiliencia es considerada como tolerancia a fallas aleatorias de los componentes de la red, comúnmente se le llama *robustez* [17]. En la literatura, los conceptos *flexibilidad* y *confiabilidad* están relacionados también [52, 71].

Los orígenes de las fallas aleatorias en los sistemas interconectados incluyen la sobrecarga de componentes, pérdida de energía o enlaces de telecomunicación, configuración errónea de los componenetes y fallas de dispositivos electrónicos.

Cuando el interés no yace en las fallas aleatorias sino en qué tan bien un sistema interconectado soporta ataques deliberados [18], el término *vulnerabilidad* es usado. El objetivo usual de aquel atacante es causar la mayor ruptura al sistema [29, 92] y el ataque puede consistir en una secuencia de ellos en vez de uno sólo [77].

Cuando múltiples amenazas existen para un sistema, la vulnerabilidad del mismo, como un todo, puede ser cuantificada en términos de los impactos individuales y en la ocurrencia de probabilidades de estas amenazas [72]. Éste se extiende a la robustez sobre fallas aleatorias cuando los impactos y las probabilidades pueden ser estimados. Entonces, los eventos posibles pueden ser categorizados desde el más grave (el de mayor impacto y probabilidad) hasta el menos grave.

En ambos casos, robustez y vulnerabilidad, el impacto del daño no es la única medida interesante, sino también la velocidad en la cual el sistema es capaz de recuperarse [53].

Las estrategias y metodologías para controlar el “ciclo de la resiliencia” (refiriéndose a los cambios de estado entre operacional, interrumpido y recuperado) [72, 73] enfatiza cómo los sistemas son propensos a fallas y que un completo entendimiento de la operación normal de un sistema es la base necesaria para la identificación de posibles problemas. Los pasos básicos involucran *identificar* amenazas, *detectar* eventos indeseables y tener formulados los *planes de respuesta* para cualquier evento que pueda comprometer la funcionalidad del sistema.

3.1.1. Teoría de grafos

Un *grafo* G es un par de conjuntos: los *vértices* (también conocidos como nodos) y las *aristas* (también conocidas como arcos o enlaces). Cada vértice $v \in V$ representa un elemento de interés, posiblemente asociado con un vector de propiedades (como el nombre, costo, capacidad, tipo), mientras que cada arista $(u, v) \in V \times V$ conecta el vértice $u \in V$ con el vértice $v \in V$, también pueden tener un vector de propiedades asociado (representando la distancia, tasa de transferencia, etc.).g

Una arista es *dirigida* si la conexión que representa es solo admitida en una dirección, desde el origen u hasta el destino v , mientras que una arista *no dirigida* representa una conexión bidireccional. Una arista es *reflexiva* si su origen y destino son el mismo vértice.

El *grado* de un vértice $d(v)$ es el número de aristas incidentes hacia él, esto es, el número de vértices que tiene v como destino. Para los grafos no dirigidos, no importa si v es el origen o el destino; para grafos dirigidos, el número de aristas con v como origen es definido como *grado de salida* $d^o(v)$, mientras que el número de aristas con v como destino es definido como *grado de entrada* $d^i(v)$. El conjunto de vértices adyacentes a v en G es denotado como $\Lambda(v)$ y llamado *vecinos*, nuevamente la distinción entre $\Lambda^o(v)$ y $\Lambda^i(v)$ existe para grafos dirigido.

La *secuencia de grados* de un grafo es un vector formado por los grados de los vértices y la *distribución de grados* se refiere a la relación entre los grados presentes en el grafo y su frecuencia. El *grado máximo* es denotado comúnmente como Δ y aquellos vértices con alto grado son llamados *hubs* (“centro de actividad” en español), sobre todo cuando la distribución de grados está lejos de ser una distribución uniforme. En redes naturales, una forma común de la distribución de grados es la llamada *ley de potencias* (power law), donde la frecuencia del grado es proporcional a la potencia negativa, $\Pr[d(v) = k] \sim k^{-\gamma}$ [2], aquellos grafos son llamados *libres de escala* (scale-free).

Típicamente, los grafos que representan redes del mundo real son esparcidos, es decir que el conjunto de vértices E es un conjunto relativamente pequeño de $V \times V$. Una medida típica para capturar tal fenómeno del grafo $G = (V, E)$, expresado en términos de $n = |V|$ (llamado el *orden* del grafo) y $m = |E|$ (llamado el *tamaño* del grafo) es el *grado medio*

$$d = m/n, \tag{3.1}$$

una variante, con el objetivo de normalizar el entre un rango específico (dependiendo si las aristas son dirigidas, reflexivas, etc) es la proporción del tamaño del grafo con el máximo tamaño posible del grafo dado el orden del mismo, para el caso no dirigido y no reflexivo la fórmula usada es

$$\delta = \frac{2m}{n(n-1)} \quad (3.2)$$

el cual equivale a 1 si todos los pares de vértices están conectados por una arista y 0 si no existen aristas en el grafo.

Un *subgrafo* S de G posee vértices como un subconjunto V_S de V y el conjunto de las aristas E_S es un subconjunto de las aristas en E , las cuales poseen nodos de origen y destino en S . Si todas estas aristas están presentes, el subgrafo es llamado *inducido*. La densidad del subgrafo inducido por $\Lambda(v)$ se le conoce como el *coeficiente de clusterización* de v , denotado por $c(v)$. El promedio global de $c(v) \forall v \in V$ es llamado coeficiente de clusterización del grafo.

Un *camino* es una secuencia de aristas que parten de un vértice v a un vértice u . Un camino es *simple* si no contiene aristas repetidas. El *largo de un camino* $\ell(v, u)$ es el número de aristas que contiene un camino simple. A veces este largo es considerado en términos del número de vértices el cual es $\ell(u, v) + 1$ si $v \neq u$, mientras que si $u = v$, el camino es llamado *ciclo*. Dado un par de vértices u y v , puede haber varios caminos que los conecten, el largo del camino más corto es llamado *distancia* $D(u, v)$. La máxima distancia entre todos los pares de vértices del grafo es llamado *diámetro* del grafo, \mathcal{D} . La distancia promedio es, evidentemente, el promedio de todas las distancias para todo par distinto de vértices. Además, la distancia de un vértice consigo mismo es considerada 0 dado que el conjunto vacío es un camino válido.

Un conjunto de vértices que están conectados entre sí dentro de un grafo es llamada *componente* (conectada). En los grafos dirigidos, una conectividad fuerte se refiere a la existencia de caminos en ambas direcciones, mientras que la k -conectividad se refiere a la existencia de múltiples caminos que no comparten aristas y/o vértices intermedios. Un grafo con solo una componente se le llama *conexo*, mientras que uno con múltiples componentes está desconectado.

Las aristas de un grafo pueden ser fácilmente expresada a través de la *matriz de adyacencia* \mathbf{A} , donde el conjunto de vértices es etiquetado como $V = \{1, 2, \dots, n\}$ de manera que se pueda asociar el elemento a_{ij} a si los vértices i y j están conectados por una arista o no:

$$a_{ij} = \begin{cases} 1, & \text{si } (i, j) \in E, \\ 0, & \text{si no.} \end{cases} \quad (3.3)$$

Donde la matriz \mathbf{D} es la matriz diagonal de grados derivada de \mathbf{A} como la suma por filas:

$$d_{ij} = \begin{cases} d(i), & \text{if } i = j, \\ 0, & \text{otherwise.} \end{cases} \quad (3.4)$$

La matriz *Laplaciana*

$$\mathcal{L} = \mathbf{I} - \mathbf{D}^{\frac{1}{2}} \mathbf{A} \mathbf{D}^{-\frac{1}{2}}, \quad (3.5)$$

donde \mathbf{I} es la matriz identidad, es una herramienta útil para estudiar las propiedades estructurales de un grafo [16], especialmente a través de su *espectro*, que corresponde a la secuencia de valores propios de \mathcal{L} , $\lambda_1, \dots, \lambda_n$, ordenados de tal manera que $\lambda_i \leq \lambda_j$ para $i < j$. La multiplicidad de λ_1 corresponde al número de componentes conexos.

3.1.2. Redes Complejas

En general, agregar *redundancia* a un sistema puede servir para aumentar su robustez [39], aunque esto evidentemente tiende a incrementar los costos de construir y mantener el sistema en cuestión. Además, la redundancia sola no garantiza robustez, es necesario poseer una distribución de grados libres de escala para que funcione [1]. En general, las distribuciones libres de escala y bimodales se han establecido como más resilientes [8, 77], incluso cuando existen diferencias en términos, bien vulnerabilidad ante ataques o robustez ante fallas aleatorias [1, 8]. Los ataques tienden a focalizarse en maximizar el daño, los grafos regulares son menos vulnerables debido a la carencia de objetivos evidentes [21]. Además, solamente permitiendo una distribución uniforme es útil si la conectividad es lo suficientemente alta a pesar de que el grafo sea esparcido [29]. También el efecto de las correlaciones de los grados (es decir, la manera en que los grados de los extremos de los enlaces están correlacionados) se ha estudiado en términos de resiliencia [49], un grafo es llamado *disasortativo* si los vértices de grados muy distintos están conectados y *asortativo* cuando los vértices de grado similar están conectados.

El esfuerzo de modificar un sistema para incrementar su resiliencia es una tarea de *optimización estructural*, donde algunas funciones objetivo que miden la resiliencia son optimizadas agregando, eliminando o modificando los vértices y las aristas, mientras que un conjunto de restricciones (como la ubicación, capacidad, costo) son respetadas [5]. Otra alternativa para los sistemas que contienen componentes “inteligentes”, que son capaces de alterar su configuración o comportamiento, es permitir que el sistema se adapte y así dinámicamente recuperarse de eventos desfavorables. Tal adaptación requiere la identificación de las posibles amenazas, generando acciones posibles para cada una y estimando su impacto, seleccionando y ejecutando una acción, y posteriormente evaluando su efecto para determinar si una recuperación adecuada ha sido alcanzada [15].

Una alternativa abordable para los sistemas adaptativos o estructuralmente optimizados es *simular* el comportamiento de un sistema de interés: identificar un peligro potencial, recrearlo en una simulación y explorar las posibles acciones para superarlo [68].

El estudio de las redes complejas ha dado lugar a un campo de investigación que propone y analiza *generadores* que artificialmente producen topologías de redes con algunas características deseadas, tales como las del Internet [51]. Los modelos clásicos son aquellos de Gilbert [36] y Erdős–Rényi [27, 28] que sitúa aristas uniformemente al azar entre los vértices, resultando

distribuciones de grado binomiales, baja distancia promedio y baja densidad local. Watts y Strogatz [84] propusieron un modelo que combina una relativa baja distancia promedio con altas densidades locales. Luego, Barabási [3] propuso un modelo donde la distribución de grado es libre de escala, con baja distancia promedio pero con densidad local baja. De aquí en adelante, cientos de modelos de generación han sido propuestos para todas las imaginables combinaciones de propiedades estructurales. Usando tales generadores para explorar las topologías posibles de los sistemas previo a construir un sistema verdadero es una manera costo efectiva para examinar la resiliencia de un sistema incluso antes de decidir cómo construirlo.

3.2. Métricas existentes

Para cualquier métrica M , *deterioration* ($DEPENDS$, $[V]$) puede ser calculada como la diferencia entre el valor de la métrica para toda la red M_0 , y el promedio del valor de la métrica en variados porcentajes de falla M_i , normalizada por M_0 [7].

3.2.1. Métricas basadas en eliminación

Muchas de las métricas basadas en eliminación de vértices son funciones de el número de vértices eliminados del grafo. Estas definiciones pueden ser aplicadas o bien para análisis del peor caso, caso promedio, o como un muestreo aleatorio repetitivo de cuanto a qué subconjunto es eliminado (la eliminación de distintos subconjuntos de la misma cardinalidad puede tener efectos drásticamente distintos en la estructura del grafo dependiendo de cuales vértices son eliminados). La eliminación aleatoria es considerada como modelo de *fallas*, mientras que la eliminación en el peor caso es considerada como modelo de un *ataque*. La métrica *connectivity robustness function* ($POLY$, $[Con, D]$) [87] es calculada eliminando (al azar) $k \in [0, n]$ vértices del grafo de entrada y calculando el orden de componente conectada más grande para cada k , normalizada por $n - k$ de manera que los valores resultantes sean dentro de $[0, 1]$.

La conectividad y el balance de carga en los sistemas interconectados sometidos a eliminaciones de vértices o aristas han sido ampliamente estudiados [5, 10, 71]. El modelo matemático del fenómeno está relacionado con la *percolación* [56].

Adicionalmente, sólo hablando sobre conectividad en términos de la existencia de caminos, métricas relacionadas con un servicio han sido propuestas: un grafo orientado al servicio S es *autosuficiente* (inglés: self-sufficient), con respecto a los requerimientos del servicio, si cada servicio requerido por $v \in V_S$ puede ser suministrado por $w \in V_S$. Un grafo G se le dice *resiliente ante k fallas* (inglés: k -failure resilient) en términos de los vértices si un subconjunto arbitrario de k o menos vértices de G pueden fallar y aun así el subgrafo resultante es autosuficiente. Con estas definiciones, en [63] se propone una métrica llamada *node resilience* (EXP , $[Con]$) como el k más grande para el cual la red es resiliente a k fallas. La métrica propuesta también puede extenderse a aristas: *edge resilience* como el entero k más grande tal que el grafo es resiliente ante k fallas de enlaces, lo que significa que el grafo resultante

es autosuficiente ante la pérdida de k o menos enlaces arbitrarios.

Una métrica definida por eliminación de subconjuntos de vértices de cardinalidad k , es *k-resilience factor* (*EXP*, [*Con*]) [65], la se construye mediante el porcentaje de grafos inducidos conectados, resultantes de la eliminación de $k - 1$ vértices de todos los posibles (conectados y desconectados) grafos inducidos. Promediando esto por $\forall k \in [2, n - 1]$ obtenemos *resilience factor*. Notar que el número de posibles subgrafos para un grafo tiene un comportamiento asintótico exponencial¹, por lo que el cálculo es computacionalmente demandante incluso para grafos de orden moderado.

Si generalizamos la eliminación de vértices a perturbaciones arbitrarias de la estructura del grafo, el cálculo puede ser formulado como la diferencia entre el orden del componente conectado más grande del grafo original con el grafo perturbado, llamada *perturbation score* por [62] (*POLY*, [*Con*]). La robustez de diferentes grafos puede ser comparada entre sí aplicando una perturbación similar a ambos y calculando la diferencia de *perturbation score* obtenida. Dados dos grafos como input, ellos definen *preferential perturbation score* (*POLY*, [*Con*]) como aquella diferencia. Luego definen *maximum perturbation score* como el valor que es el máximo de *preferential perturbation score*, alcanzado por la perturbación que maximiza el daño en el primer grafo y minimiza el daño en el segundo.

Otro método que toma en cuenta todos los vértices en vez de solo los que pertenecen al componente conectado más grande es *pairwise disconnectivity index* [61] (*POLY*, [*Con*]), el cual es calculado en términos de los pares de vértices que están conectados en el grafo original $\mathcal{P}(V)$: *pairwise disconnectivity index* de un vértice v es el número de pares de vértices que siguen conectados luego de remover v del grafo $\mathcal{P}(V \setminus \{v\})$, que además puede ser naturalmente generalizado para la eliminación de conjunto de vértices en vez de uno solo.

En [69] se propone la siguiente métrica, llamada *unique robustness measure* (*POLY*, [*Con*]), para los ataques basados en grados donde los vértices son eliminados de mayor a menor grado (con o sin recalculando los grados después de la eliminación de los vértices) y después de cada eliminación, el orden del componente conectado más grande C_t es calculado. La suma de C_t sobre $t = 1, \dots, n$ es luego normalizada por n y usada como una métrica de robustez.

3.2.2. Métricas basadas en caminos

Una forma común para caracterizar la estructura de un grafo es calculando las distancias de todos los pares de vértices. La *eficiencia* de $u, v \in V$, $u \neq v$, es $\varepsilon_{u,v} = 1/D(u, v)$; en [44] miden la tolerancia a fallas en términos de la *eficiencia local*. Luego en [87] definen *efficiency function* (*POLY*, [*Con*, *D*]) como el promedio de la eficiencia para todos los pares de nodos, normalizado por el la cantidad de pares de nodos del grafo.

Otra, muy popular, forma de medir basándose en caminos es la *centralidad* (inglés: *betweenness*) refiere al número de caminos más cortos que atraviesan una arista (o vértice), ponderado de tal forma que si vértices v y u se conectan con k caminos de largo mínimo,

¹Los posibles subconjuntos de $n - k$ vértices es el coeficiente binomial para n y k , el cual varía desde $k \in [0, n]$ sumando 2^n , y por ende para el rango propuesto, la suma es $2^n - 2n - 1 \in \mathcal{O}(2^n)$.

cada uno de ellos cuenta como $1/k$ en la centralidad [55]; también se puede estimar con caminatas aleatorias [54]. Posteriormente en [66] se define *relative betweenness* ($POLY, [Cen]$), obtenida normalizando la centralidad de un vértice por la suma de todos los valores de centralidad de todos los vértices del grafo. Luego después calculan *betweenness interference* ($POLY, [Cen]$) como la diferencia de *relative betweenness* de un vértice v antes y después de la eliminación de otro vértice w .

En [67] proponen una métrica llamada *temporal efficiency* ($POLY, [D]$), definida en términos de la manera en la cual la eficiencia (anteriormente definida) de un grafo se degrada (como un porcentaje de la eficiencia original) después de algún daño hecho (como la eliminación de un vértice o arista, o bajando el nivel de funcionalidad de un vértice o arista si el grafo modela algún sistema en que los elementos pueden estar parcialmente operativos)

El autor en [59] denota el estado de percolación de un nodo i en un tiempo t como x_i^t . Cuando el contexto temporal esta claro, simplemente usamos x_i . Específicamente, $x_i^t = 0$ indica un estado no percolado y $x_i^t = 1$ indica un estado completamente percolado, mientras que un estado parcialmente percolado corresponde a $0 < x_i^t < 1$.

La fracción

$$w_{s,v}^t = \frac{x_s^t}{[\sum x_i^t] - x_v^t}, \quad (3.6)$$

muestra que tanto del estado de percolación total es el estado del nodo s , excluyendo v . En [59] se define un *camino percolado* como el camino más corto de $v \in V$ a $w \in V$ tal que v esta infectado, y luego define *percolation centrality* ($POLY, [P, Cen]$) como la proporción de caminos percolados que pasan a través él sobre todos (percolados o no) caminos:

$$PC^t(v) = \frac{1}{N-2} \sum_{s \neq v \neq r} \frac{\sigma_{s,r}(v)}{\sigma_{s,r}} \frac{x_s^t}{[\sum x_i^t] - x_v^t}, \quad (3.7)$$

$$PC^t(v) = \frac{1}{N-2} \sum_{s \neq v \neq r} \frac{\sigma_{s,r}(v)}{\sigma_{s,r}} w_{s,v}^t. \quad (3.8)$$

Si cada nodo tiene el mismo nivel de percolación > 0 , $w_{s,v}^t$ equivale a $\frac{1}{N-1}$, luego $PC^t(v) = BC(v)$, donde $BC(v)$ es la centralidad de ese nodo.

Otra métrica relacionada con caminos, es *global functionality loss* ($POLY, [D]$), definida por [91] en términos de la funcionalidad de un nodo ante un ataque I (el cual remueve un nodo w):

$$F_I(v) = F_{I-1}(v) - \frac{1}{D(v,w)^2 d(v)} F_{I-1}(v). \quad (3.9)$$

Donde $D(v, w)$ y $d(v)$ corresponden a la distancia y el grado después del ataque $I - 1$. Esta definición se construye fijando todas las funcionalidades iniciales como funcionalidades no afectadas, es decir $F_0(v) = 1, \forall v \in V$. Usando esta funcionalidad, y asumiendo que en cada ataque se elimina un vértice, se define la pérdida de la funcionalidad como la suma de las diferencias de la funcionalidades de un nodo antes y después del ataque para cada ataque sobre esta secuencia de k ataques (esto también permite extender para la eliminación de enlaces en vez de nodos):

$$FL(v) = \sum_{I=1}^k F_{I-1}(v) - F_I(v). \quad (3.10)$$

Así posteriormente *global functionality loss* (FLN) se obtiene siguiendo la siguiente fórmula:

$$FLN = \sum_{v, v \neq w_I}^N FL(v), \quad (3.11)$$

donde w_I corresponde al nodo eliminado por el ataque I .

3.3. Aplicaciones

3.3.1. Transporte

Cuando el sistema bajo estudio *transporta* algo (bien puede ser algo físico o tráfico de redes), no es suficiente modelar y caracterizar la estructura del sistema sino también las dinámicas de carga que sí misma necesitan ser tomadas en cuenta.

En una topología de “rutas activas” se sobrepone con las decisiones de ruteo encima de la topología de la red en sí. La forma en que se dirige el tráfico a las rutas disponibles es en esencia un problema de optimización multicriterio [6, 79]. Lo mismo aplica a tráfico vehicular en cuestiones como control de semáforos, recomendación de rutas y vehículos autónomos.

En las redes de transporte, los enlaces representan las relaciones (camino, segmentos de ferrocarril, vuelos, etc.), y cada uno posee un peso asociado llamado el *tiempo de viaje* (esto puede ser determinístico o estocástico, dependiendo del modelo), y la cantidad de tráfico en un enlace es llamada *flujo* sobre tal enlace. Cuando los flujos son conocidos, *vehicle-hours of travel* (VHT) (horas-vehículo de viaje) puede ser calculado como la suma ponderada de los flujos sobre los enlaces, usando los tiempos de viaje como pesos, en equilibrio².

²Una red de transporte está en *equilibrio* cuando no es posible para ninguno de sus usuarios reducir sus tiempos de viajes con un cambio unilateral de ruta (cónfer a game-theoretical equilibrium [83]).

Ahora, si un enlace es agregado o eliminado, el VHT cambiará, y la magnitud de tal cambio puede ser usada como métrica para el impacto de tal enlace, notar que eliminando un enlace no solo afecta el flujo en el mismo si no que también puede forzar cambios de flujo en otros enlaces a fin de que un equilibrio sea alcanzado. En [57] se propone una métrica llamada *network trip robustness* (*POLY*, [D]), calculada como la suma de cambios de VHT sobre todos los enlaces, normalizada por la demanda total del grafo,

Otra métrica aplicada en este contexto utiliza los conceptos de distribución de grado y su entropía relativa.

La entropía es una medida de incertidumbre de variables aleatorias, y también es una medida de cantidad de información requerida para describir una variable aleatoria en un sentido común.

La entropía relativa (o distancia Kullback-Leibler) es una medida de distancia entre dos distribuciones aleatorias. En estadística, corresponde al logaritmo de la esperanza de los radios de probabilidad. Toma dos funciones de densidad de probabilidad $p(x)$ y $q(x)$, y se define como sigue:

$$RE\left(\frac{p}{q}\right) = \sum_{x \in A} p(x) \log\left(\frac{p(x)}{q(x)}\right). \quad (3.12)$$

Típicamente, $p(x)$ representa la distribución verdadera, mientras que $q(x)$ representa la distribución teórica. En [33] se propone que $p(x)$ sea considerada como la distribución de grado de la red de ferrocarril a estudiar, y $q(x)$ sea tratada como la red de ferrocarril con una distribución completamente desordenada.

Luego, usando la entropía relativa para medir la robustez, se define $q(x)$ como una distribución de grado uniforme $1/N$ (red desordenada), y $p(x)$ como la verdadera distribución de grado. Así la ecuación 3.12, la cual define *relative entropy* (*ND*, [Deg]) se vuelve:

$$RE(p) = \sum_{x \in V} p(x) \log(Np(x)). \quad (3.13)$$

En las redes de tráfico aéreo, bajo el contexto de el problema de máximo flujo³, se define en [58] la métrica *relative area index* (*ND*, [Cen]) de un nodo en función de su parámetro de degradación u y el flujo máximo del nodo $F_{max}^v(u)$. El parámetro de degradación representa la reducción de la capacidad de vuelo de algún nodo en específico. Se consideran dos tipos de degradación: (1) basadas en porcentaje; (2) basadas en valor absoluto. La basada en porcentaje varía entre 0% y 100% reducción de la capacidad del nodo. Mientras que la absoluta varia entre 0 y algún valor predeterminado α (medida en vuelos por hora). Así asumiendo que $F_{max}^v(u)$ es monótonicamente decreciente:

³El problema de máximo flujo involucra la búsqueda de un flujo factible a través de una red de flujo de una sola fuente y un solo sumidero tal que es máxima [70].

$$RAI(v) = \frac{\int_{u_0}^{u_T} w(u)(F_{max}^v(u_0) - F_{max}^v(u))du}{\int_{u_0}^{u_T} w(u)F_{max}^v(u_0)du}, \quad (3.14)$$

donde u_0 y u_T son los valores mínimos y máximos permitidos para el parámetro de degradación, es decir $[0, 1]$ y $[0, \alpha]$ respectivamente para el valor en porcentaje y en absoluto. Notar que dos funciones de flujos distintas pueden poseer la misma área, El valor $w(u)$ corresponde al parámetro de peso que se le asigna para distinguir estas funciones, asignando diferentes prioridades a los diferentes rangos de reducción de capacidad, dependiendo del tipo del nodo de interés, su naturaleza y capacidad de reducción y escenarios de aplicación. La intuición detrás de esta métrica es darle más importancia al nodo si éste se degrada más rápido “en promedio” (ya que no es una suma, sino una integral), donde el denominador corresponde a una normalización del valor para comparar diferentes nodos.

3.3.2. Red Eléctrica

Las redes eléctricas son indispensables hoy en día. Existe una gran dependencia en los países desarrollados con la energía eléctrica y por ende es importante contar con que la red sea confiable y disponible todo el tiempo. Los apagones pueden ser muy dañinos y pueden paralizar la industria de casi cualquier índole. Otras investigaciones indican que los apagones son normalmente causados por fallas sucesivas de una gran porción de componentes desencadenadas por una perturbación inicial (Por ejemplo, rayos, desastres naturales, contacto con las vegetaciones o error humano).

Avances recientes en las redes complejas muestran un potencial prometedor para evitar estas fallas, detectando, evaluando y mejorando la robustez del sistema.

Bajo este contexto, en [42] se propone la métrica *network robustness metric* (*POLY*, [*Deg*]) que corresponde a la suma de la significancia y robustez nodal de cada nodo.

La robustez nodal se refiere, dado un nodo, a la suma de los valores del flujo normalizado (medido en Watts (W)) de cada enlace conectado a ese nodo, ponderado por el logaritmo de este flujo normalizado junto con su parámetro de tolerancia α_i de ese enlace, es decir,

$$R_{v,i} = - \sum_{i=1}^{L_v} \alpha_i p_i \log(p_i), \quad (3.15)$$

donde L_v es el tamaño del conjunto de enlaces conectados al nodo v . El signo menos $-$ es usado para compensar el valor negativo que ocurre por el logaritmo de los valores de flujo normalizados,

$$p_i = \frac{f_i}{\sum_{j=1}^E f_j}, \quad (3.16)$$

(los logaritmos en $(0, 1)$ corresponden a valores negativos). El parámetro de tolerancia de un enlace es simplemente el inverso del nivel de carga de un enlace arbitrario i , el cual es la proporción entre la carga y la capacidad máxima de aquel enlace. La carga es frecuentemente medida como la centralidad del aquel enlace y la capacidad medida como el máximo flujo de energía que puede pasar por el enlace.

La significancia nodal es la potencia distribuida por el nodo, normalizada por la suma de todas las potencias de todos los nodos.

Otra medida de robustez, esta vez global, propuesta en [43], es *effective graph resistance* (*POLY*, [Con]), que corresponde a la suma de todas las resistencias efectivas entre todos los pares de nodos de la red. La resistencia efectiva es una medida de “distancia de camino eléctrico” entre dos nodos. Utiliza la matriz Laplaciana \mathcal{L} , específicamente, en grafos ponderados:

$$\mathcal{L}_{ij} = \begin{cases} d(i), & \text{si } i = j \\ -w_{ij}, & \text{si } i \neq j \text{ e } (i, j) \in E \\ 0, & \text{si no.} \end{cases} \quad (3.17)$$

Considerando w_{ij} es la ponderación del enlace (i, j) . Así, la resistencia efectiva del link (i, j) es medida como:

$$R_{ij} = \mathcal{L}_{ii}^+ - 2\mathcal{L}_{ij}^+ + \mathcal{L}_{jj}^+, \quad (3.18)$$

donde M^+ el pseudo-inverso de Moore-Penrose de la matriz M^4 .

3.3.3. Telecomunicaciones

La comunicación a través de sistemas electromagnéticos puede ser fácilmente modelada a través de las redes complejas, y, naturalmente medir su robustez asegura que aquella comunicación sea segura y confiable. Internet es un gran ejemplo, la cual es casi imprescindible hoy en día. Crear una red resiliente nos permite comunicarnos en caso de desastres o peligros, lo cual permite una clara y rápida respuesta ante aquellos problemas.

Una forma de modelar el comportamiento de un usuario de Internet, es con caminatas aleatorias. Un usuario estando en un sitio, correspondiente a un nodo s en la red navega impredeciblemente entre los enlaces de las páginas hasta finalizar su “viaje” en su destino final d . Las páginas las cuales visitó, y, los enlaces por los cuales se dirigió corresponden a un camino en sobre la red subyacente.

Así, para cuantificar que tanto cambia la capacidad de la red ante la eliminación de un nodo, en [78] definen *node criticality* (*POLY*, [Cen]). Ésta corresponde a la centralidad basada en caminos aleatorios del nodo ponderada por su peso. La centralidad basada en caminos

⁴Corresponde a una generalización del inverso de una matriz [25].

aleatorios (inglés: random-walk betweenness) $b_{sk}(d)$ de un nodo k dados un par origen—destino (s, d) , se refiere al número esperado de veces que una caminata aleatoria pasa por k en su viaje desde el origen s al destino d . La centralidad total del nodo k es la suma de esta cantidad para todo los pares (s, d) posibles. Naturalmente, el autor extiende esta definición para los enlaces también.

Luego, usando esta métrica, el autor crea una métrica global de robustez *network criticality* (*POLY*, [*Cen*]), que simplemente es la suma de *node criticality* de cada nodo.

Otra opción estudiada para modelar robustez en redes de telecomunicaciones, corresponde al flujo de datos. Un flujo de datos (de ahora en adelante simplemente llamada flujo) es una asociación de datos entre un par de nodos que puede ser distribuido sobre uno o más caminos. Un flujo es considerado confiable si al menos un camino sigue conectado durante una falla (interrupción o eliminación de algún nodo o enlace de la red). Así nace la robustez de los flujos, que simplemente, es la cantidad de flujos confiables dividida por la cantidad total de flujos de la red.

Luego en [14] definen *aggregated remaining flow* (*POLY*, [*Con*]) como el promedio de la robustez de flujos de la red, después de cada ataque (eliminación de un nodo), dentro de un conjunto de ataques predeterminado. El autor extiende esta métrica para comparar diferentes grafos, llamada *normalized aggregated remaining flow* (*POLY*, [*Con*]), normalizando por la cantidad de enlaces del grafo.

En el mismo estudio, también se utiliza la diversidad de caminos [73] que toma como medida el camino óptimo P_0 y otro camino P y calcula:

$$D(P) = 1 - \frac{|P \cap P_0|}{|P_0|}, \quad (3.19)$$

donde los conjuntos se pueden manejar o a base de los vértices o a base de las aristas.

Usando la diversidad de caminos, se puede extender para un conjunto de caminos, diversidad de caminos efectiva, definido como la utilidad (ponderada) de un conjunto caminos alternativos simples,

$$EGPD = 1 - e^{-\lambda k_{sd}}. \quad (3.20)$$

Donde λ es una constante de peso para la utilizada de cada camino adicional añadido y k_{sd} es la suma de las diversidades de cada camino:

$$k_{sd} = \sum_{i=1}^m D(P_i). \quad (3.21)$$

Así, para todos los caminos se define la diversidad de caminos total como el promedio de los diversidades de caminos efectiva para cada par de nodos del grafo.

En el mismo estudio [14], se redefine la diversidad de caminos como *geographical path diversity* (*POLY*, [*Con*]) $D(P_a)$, como la distancia geográfica mínima $D \geq d$ entre cualquiera de los nodos del camino P_a con los del camino más corto (excluyendo los nodos de origen y destino).

Naturalmente, extienden todas las definiciones anteriores en el sentido geográfico, esto es, diversidad de caminos efectiva a *effective geographical path diversity* (*POLY*, [*Con*]) y diversidad de caminos total a *total graph geographical diversity* (*POLY*, [*Con*]). Finalmente se puede calcular *compensated total geographical graph diversity* (*POLY*, [*Con*]) para comparar un grafo G con el más grande grafo M (en cantidad de enlaces) dada la topología en consideración:

$$cTGGD = e^{TGGD-1} \left(\frac{|E(M)|}{|E(G)|} \right)^{-\rho}, \quad (3.22)$$

donde ρ es una constante experimentalmente elegida (y para términos prácticos puede omitirse).

3.3.4. Gestión de cadenas de suministro

Cuando el sistema consiste en la producción, transporte y consumo de algún elemento podemos hablar de redes de suministro. Una red de suministro usualmente consiste en productores, nodos de suministro que desarrollan un bien, y consumidores, nodos de demanda que requieren de tal bien para que el sistema sea factible.

Un caso especial de *connectivity robustness function*, para el caso de las redes de suministro, se aplica redefiniendo el componente conectado como uno donde no sólo la conectividad es requerida, sino también requiere que al menos un nodo de suministro debe estar presente en cada componente (aquellos componentes sin nodos de suministro son ignorados en el cálculo), esto es llamado *network connectivity* en [93] (*POLY*, [*Con*, *D*]). Es decir, se define como el tamaño del componente más grande que posea un nodo de suministro.

Bajo el mismo contexto, el autor también define *supply availability* (*POLY*, [*Con*, *D*]), que corresponde al porcentaje de nodos de demanda que tienen acceso (es decir, que existe un camino) a, al menos, un nodo de suministro.

Usando nuevamente el concepto de eficiencia, el autor plantea dos métricas más *best delivery efficiency* (*POLY*, [*Con*, *D*]) y *average delivery efficiency* (*POLY*, [*Con*, *D*]). La primera corresponde al recíproco del promedio, para cada nodo de demanda, de el camino más corto hacia algún nodo de suministro. La segunda corresponde al promedio de las eficiencias promedio hacia cada nodo de demanda (donde cada camino posee hacia un nodos de suministro posee una ponderación $f(u)$). Específicamente, para cada nodo la eficiencia promedio corresponde:

$$ADE(v) = \sum_{u=1}^{|K(v)|} (\varepsilon_{u,v})^{1/f(u)}, \quad (3.23)$$

donde $\varepsilon_{u,v}$ corresponde a la eficiencia entre el nodo u y v . La función $f(u)$ puede ser omitida y describe la importancia de cada camino hacia el nodo de suministro u . Así, se calcula el promedio y posteriormente se normaliza y obtenemos:

$$ADE(G) = \frac{1}{|V_D|} \sum_v^{|V_D|} ADE(v). \quad (3.24)$$

Adicionalmente, con el modelo de cadenas de suministro, se puede diferenciar la cantidad del producto recibido en cuestión o también un nodo de demanda puede recibir cantidades independientemente diferentes de suministro de cada nodo que lo proporcione. Una métrica que utiliza esta idea, llamada *supply chain index* (*POLY*, [*Cen*]), fue propuesta en [60] y diseñada para señalar la importancia de cada nodo. Primero, se calcula s_{ji} que representa la proporción del producto total recibido por el nodo j del nodo de suministro i relativo al total de producto recibido (y quizás eventualmente transferido) por j . Esto cumple con que para un nodo cualquiera j , $\sum_i s_{ji} = 1$. Luego, una segunda variable p_j , que mide la proporción del producto total de la cadena de suministro para un nodo j . La métrica para cada elemento j se define de la siguiente manera:

$$SCI_j = \sum_{i=1}^n s_{ji} p_j^2. \quad (3.25)$$

3.3.5. Biología y medicina

En la biología existen diversos métodos para estudiar los fenómenos químicos e interacciones. La teoría de grafos es ampliamente usada para caracterizar diversas relaciones entre distintos elementos biológicos. Por ejemplo, las múltiples reacciones químicas pueden ser modeladas por un grafo, donde cada nodo puede ser un compuesto químico (como una proteína) y cada enlace una reacción química, medir la robustez en este contexto puede indicar que proteína puede ser vital en una cadena de reacciones como otras que no, o bien puede indicar que tan frágil es esta cadena de reacciones ante la privación de alguna enzima. Estas redes suelen ser extensas y un gran número de ellas tienden a seguir distribuciones de grado con leyes de potencia.

Una métrica que se usa en un contexto similar corresponde a *resilience score* la cual se divide en dos estrategias *random resilience score* (*ND*, [*Con*]) y *targeted resilience score* (*ND*, [*Con*]) propuestas en [75]. Cada nodo del grafo representa un metabolito⁵ y cada enlace representa una reacción química entre metabolitos. Las métricas propuestas corresponde a

⁵Cualquier sustancia producida durante el metabolismo (digestión u otros procesos químicos corporales).

la proporción de número de metabolitos a remover (al azar para *random* y siguiendo alguna regla de ataque para *targeted*) hasta que la ley de potencia (de la distribución de grado) se rompa, sobre la cantidad total de metabolitos.

A veces, en vez de proteger una red queremos destruirla, pero además, como exigencia adicional se requiere preservar otra red. Tal es el caso de las redes de células cáncer, donde el objetivo es causar el mayor daño a las células malignas mientras que se trata de minimizar el daño en aquellas benignas. Usando el concepto de perturbación, el mismo autor [62] define *average perturbation gradient* (*POLY*, [Con]), que corresponde al *maximum perturbation score* dividido por la proporción de el número de vértices eliminados sobre el total. Es decir,

$$APG(G_1, G_2) = \left(\frac{PS(G_1, G_2)}{N_{rm}/N} \right)_{max}. \quad (3.26)$$

La noción de esta definición no es solo maximizar el daño en una red y minimizar el daño en otra, sino que también queremos minimizar la cantidad de nodos totales eliminados.

3.3.6. Redes sociales

Algunas redes son construidas para modelar fenómenos sociales. Un ejemplo muy usado son las relaciones de artículos académicos, donde cada autor es un nodo y cada enlace es creado cuando es citado o cita a otro autor. En general, en estas redes cualquier tipo de interacción humana puede ser considerada como enlace. Medir la robustez en estas redes puede usarse para evaluar la fortaleza de una comunidad en sí o bien la importancia de algún individuo en ella.

Una métrica para medir la importancia de aquellos enlaces es *bridgeness* (*EXP*, [Con]) [13], que corresponde simplemente a la media geométrica⁶ entre el tamaño de ambos cliques máximos⁷, los cuales incluyen el nodo de origen y destino del enlace a medir, dividida por el tamaño del clique máximo que contenga el enlace.

Otra métrica utilizada para medir la robustez en redes sociales corresponde a *fragmentation* (*POLY*, [Con]) [12], definida como la proporción de pares de nodos que fueron desconectados luego de la eliminación de una fracción de nodos q del total, sobre número total de nodos de la red original.

Usando las métricas de centralidad en [19] se propone *sensitivity* (*DEPENDS*, [Deg, Cen]). Dados n variables, t_1, \dots, t_n , representando el grado de cada nodo, y dada una función de centralidad C , tal que $C(t_1)$ es la centralidad para el nodo 1 y así, *sensitivity* de el nodo i con respecto al nodo j , es el j -ésimo valor del vector de la derivada parcial con respecto a la variable j -ésima de la función C .

⁶La raíz n -ésima del producto de todos los números.

⁷Un clique es un conjunto de vértices V , tal que el subgrafo inducido por V es un grafo completo. Un clique máximo es aquel conjunto más grande que satisfaga alguna condición.

Existen casos extremos en que el grafo resultante es tan grande que las métricas convencionales no pueden ser calculadas en un tiempo razonable. Para lidiar con esto, en [50] se propone *generalized robustness index* (*POLY*, [*Cen*]), usada para aproximar el valor de otra métrica de interés llamada *normalized subgraph centrality*. Para cada nodo i en el grafo, dado un par valor propio y vector propio⁸ (λ_i, u_i) , NSC se puede aproximar como:

$$NSC_k(i) = \sum_{j=1}^k u_{ij}^2 \sinh(\lambda_j), \quad \forall i \in V. \quad (3.27)$$

Donde k es el número de valores propios que contribuirán a la aproximación de NSC . Basados en esta definición *generalized robustness index* de un grafo se define como

$$r_k = \left(\frac{1}{|V|} \sum_{i=1}^{|V|} \left\{ \log(u_{i1}) - \left(\log(\sinh^{-1/2}(\lambda_1)) + \frac{1}{2} \log(NSC_k(i)) \right) \right\}^2 \right)^{\frac{1}{2}} \quad (3.28)$$

3.3.7. Redes complejas generales

Existen métricas que, en el contexto donde fueron propuestas, no fueron aplicadas a temas específicos. Sin embargo, aquellas poseen valor científico porque se tratan de describir fenómenos abstractos que ocurren en las redes complejas y sus diversos modelos.

En la matemática existe la teoría espectral de grafos, que consiste en el estudio de las propiedades de un grafo en relación con el polinomio característico⁹, valores propios y vectores propios asociados al grafo, tanto como de la matriz de adyacencia como la Laplaciana.

Una métrica clásica, y muy estudiada, corresponde a la conectividad algebraica [34], que corresponde al segundo valor propio más pequeño de la matriz Laplaciana. Si bien la definición puede parecer arbitraria, existen muchas implicancias conocidas acerca de los valores propios de la matriz Laplaciana. Por ejemplo la cantidad valores propios equivalentes a 0, corresponde a la cantidad de componentes del grafo. Por ende, si la conectividad algebraica es mayor a 0, implica que el grafo está conectado.

Otra métrica popular, es el índice de Estrada. Fue definido por Ernesto Estrada para medir el plegamiento de proteínas [31], corresponde a la suma de las funciones exponenciales de cada valor propio de la matriz de adyacencia. Específicamente:

$$EE(G) = \sum_n e^{\lambda_j}, \quad (3.29)$$

⁸Los vectores propios de una matriz son los vectores no nulos que, cuando son transformados por el operador, dan lugar a un múltiplo escalar de sí mismos, con lo que no cambian su dirección. Aquel múltiplo escalar recibe el nombre de valor propio de la matriz.

⁹Polinomio que contiene una gran cantidad de información sobre la matriz, los más significativos son los valores propios, su determinante y su traza.

donde λ_j es el j -ésimo valor propio de la matriz de adyacencia del grafo.

Luego, una métrica que basada en el índice de Estrada, llamada *natural connectivity* (*POLY*, [*Sp*]) [85], consiste en el logaritmo natural de este índice dividido por el orden del grafo. Prontamente, en [88] se extiende a una versión local, llamada *local natural connectivity* (*POLY*, [*Sp*]), que consiste en la misma definición pero solo contando los nodos del componente conectado a medir.

Otra métrica relacionada con los valores propios fue definida en [46], llamada *reconstructibility coefficient* (*ND*, [*Sp*]). Ésta se define a la cantidad de valores propios que pueden ser igualados a cero, tal que la matriz de adyacencia pueda ser exactamente reconstruida. Una reconstructibilidad alta refleja un grafo “geoméricamente simple”, que necesita pocas dimensiones ortogonales para ser descrito.

Las distribuciones de grado también pueden ser un método para evaluar la robustez de un grafo. Así se describe en [81], donde definen *entropy of the degree distribution* (*POLY*, [*Deg*]) como una métrica de robustez. Su cálculo, autoexplicativo, corresponde a calcular la entropía de la distribución de grados de el grafo. Más entropía se traduce a más heterogeneidad, que significa mayor robustez. Notar que la entropía mínima 0 se consigue cuando todos los nodos están conectados a un nodo central singular, mientras que la entropía máxima $\log(n - 1)$ se puede lograr con una distribución aleatoria uniforme.

Por otra parte, bajo el gran crecimiento de páginas web nace la necesidad de evaluar su importancia. Modelando la estructura de Internet usando teoría de grafos, llevó a uno de los fundadores de *Google*¹⁰ a crear el indicador *pagerank* [9]. Este indicador determina la “importancia” de las páginas web basada en la “importancia” de sus páginas padres. Un algoritmo relacionando llamado *HITS* [41] también fue desarrollado por *IBM*¹¹.

Con el mismo objetivo en [22] se definen dos métricas. La primera, *entropy rank* (*ND*, [*Cen*, *Sp*]) que se basa en que cada camino que realice un navegante es igual de probable que cualquier otro. Se define, para en nodo i (de un grafo acíclico fuertemente conectado) como la probabilidad $u_i(v_i)$, donde $u(v)$ es el vector propio dominante¹² izquierdo¹³ (derecho).

La segunda, *free energy rank* (*ND*, [*Cen*, *Sp*]), similar a la definición anterior, pero reemplazando la matriz de adyacencia por otra matriz B obtenida de la matriz de adyacencia reemplazando los valores en cero por e^ε , con $\varepsilon > 0$. Ésto es para extender la métrica a grafos que no sean necesariamente fuertemente conectados o acíclicos.

Volviendo a los temas de propagación de epidemias y percolación, existe un modelo de propagación llamado redes con epidemias susceptible-infectado-susceptible (SIS). Durante la propagación, un nodo puede estar en uno de los dos estados. Primero un nodo infectado puede infectar vecinos susceptibles a un ritmo de infección β . Además puede curarse a sí mismo (y volver ser susceptible a infección) a un ritmo de curación δ . Adicionalmente, la proporción

¹⁰<https://google.com>

¹¹<https://www.ibm.com/>

¹²Aquel vector propio que posee el mayor valor propio relacionado.

¹³Vector fila que satisface la propiedad de un vector propio multiplicando a la matriz por la izquierda. La versión por la derecha, que usa un vector columna es la más usada.

entre β y δ es llamada ritmo de infección efectiva $\tau = \beta/\delta$, y su recíproco s ritmo de curación efectiva. En [89] denotan el número de nodos infectados en un tiempo t como $Y(t)$. Si la población n es lo suficientemente grande, podemos convertir $Y(t)$ a $y(t) = Y(t)/N$, lo que produce una cantidad continua que representa la fracción de nodos infectados. La métrica *viral conductance* (ND , $[P]$) de una red G con respecto a la propagación de epidemias. Representa la fracción promedio de nodos infectados para todos los tipos de ataques epidémicos capaces de producir brotes en las redes. Se define como

$$VC = \int_0^\rho y_\infty(s) ds = \rho y_\infty, \quad (3.30)$$

donde y_∞ es la función que satisface la evolución en el tiempo del epidemia SIS, ρ es el radio espectral (máximo valor propio) de la matriz de adyacencia de la red (además $y_\infty(\rho) = 0$) y y_∞ es el valor promedio de la fracción de nodos infectados para todo s , con $0 \leq s \leq \rho$.

En temas más generales de redes complejas se han desarrollado las siguientes métricas. Primero *node load* ($POLY$, $[Deg]$) [82], métrica que mide la robustez de un nodo y corresponde al grado del nodo multiplicado por la suma de los grados de los nodos vecinos, todo elevado a un parámetro variable.

Cuando se pueden modelar las probabilidades de falla de los nodos, en [45], usando el concepto de eficiencia se propone *fragility* ($POLY$, $[D]$), que corresponde al promedio de las delta eficiencias para cada nodo. La delta eficiencia de un nodo es la diferencia de los promedios de la eficiencia del grafo antes y después de remover ese nodo.

El mismo autor también propone *dynamic robustness metric* ($POLY$, $[D]$), que es el promedio de las delta eficiencia para cada nodo, pero ponderado por la probabilidad de falla de cada nodo.

Relacionado con la centralidad en [86] se calcula *subgraph centrality* (ND , $[Cen]$) de un nodo como la suma (a través de todos los largos de camino hacia infinito) de todas caminatas hacia todos los nodos de la red. Para evitar valores infinitos, el valor de la suma se divide por el factorial de largo del camino, y luego se le aplica el logaritmo natural.

$$S_C = \sum_{k=1}^{\infty} \sum_{i=1}^n (A^k)_{ii} = \sum_{k=1}^{\infty} \sum_{i=1}^n \lambda_i^k. \quad (3.31)$$

Usando series de Taylor¹⁴ $e^x = \sum_{k=0}^{\infty} x^k/k!$ y la ecuación 3.31

$$S_C = \ln \left(\sum_{i=1}^n \sum_{k=1}^{\infty} \frac{\lambda_i^k}{k!} \right) \approx \left(\sum_{i=1}^n \sum_{k=0}^{\infty} \frac{\lambda_i^k}{k!} \right) = \left(\sum_{i=1}^n e^{\lambda_i} \right). \quad (3.32)$$

¹⁴Una serie de Taylor es una aproximación de funciones mediante una serie de potencias o suma de potencias enteras de polinomios.

Una métrica simple definida en el contexto de ataque malicioso es *link robustness index* (*POLY*, [*Con*]) [90], se calcula como la fracción promedio de enlaces en el componente conectado más grande sobre el tamaño del grafo, después de $1 \dots |E|$ eliminaciones de enlaces, consecuente de un ataque.

Similarmente, otra métrica propuesta en [76] llamada *random-robustness index* (*EXP*, [*Con*]) indica la suma (por cada ronda de ataques) de la suma de el tamaño de el componente conectado más grande después de la eliminación de cada posible fracción de nodos dividida por el orden del grafo.

Particularmente, en [80] se desarrolla una métrica sin nombre, la cual llamaremos *robustness measure 1* (*POLY*, [*Con*]), la cual corresponde, dado un entero n , al número de enlaces que deben ser eliminados de la red para obtener una red con n componentes, normalizada por el orden del grafo.

Otra métrica que se basa en conjunto de ataques es *community robustness* (*POLY*, [*Con*]) propuesta en [47], se calcula dados a ataques y k comunidades, el promedio sobre todos los ataques del promedio sobre todas las comunidades de todos los nodos restantes de la comunidad (después de el q -ésimo ataque) dividido en el número total de nodos en tal comunidad. Una comunidad es simplemente un conjunto predeterminado de nodos.

Por último, una métrica relacionada con el problema de cobertura de vértices¹⁵ son propuestas en [64]. La primera, *covering degree* (*EXP*, [*Deg*]) medida a nivel de un vértice v , corresponde la cantidad de coberturas de vértices minimales que contienen aquel vértice v . La segunda *covering index*, también a nivel de un vértice v , se calcula como el número de coberturas de vértices mínimas que contienen al vértice v más la proporción de el número de coberturas de vértices minimales que contienen v sobre la cantidad de total de coberturas de vértices minimales en G .

¹⁵Una cobertura de vértices para un grafo G es un conjunto de vértices V en los que cada arco de G incide al menos en un nodo de V . La cobertura de vértices minimal es aquella que no contiene otra cobertura de vértices. La cobertura de vértices mínima es la más pequeña de las coberturas de vértices.

Capítulo 4

Caso de Estudio

En este capítulo se describirá el caso de estudio del backbone de la red física chilena. Debido a que, en el momento de la experimentación los datos no estaban completos, se decidió trabajar con una red conocida pero a menor escala. Para alcanzar un nivel mínimo de robustez, se tomó como objetivo hacer que ésta no se desconecte caso de una falla singular y aleatoria de un nodo o arista. Se propone una mejora y se evalúa un conjunto de métricas previamente seleccionadas. Debido a la poca robustez de la red, ninguna métrica corresponde a aquellas del mapeo sistemático. Se favoreció el uso de métricas básicas de robustez. Posteriormente, en base a una segunda mejora se compararon aquellas métricas y se discuten sus correlaciones.

4.1. Backbone red Chilena

Dado que no se cuenta con los datos exactos de la red física chilena, éste análisis se realiza sobre una red conocida y similar, debido a que conecta elementos a través del país. Aunque esta red es topológicamente más pequeña (contiene menos nodos y aristas), de tal manera de predecir los fenómenos que ocurrirán en la red real. La red se presenta en la figura 4.1.

Como fue mencionado anteriormente, debido a la geografía del país es complejo tener redundancia en la red. Con la existencia de **al menos un** elemento cuya eliminación desconecta la red es imposible hablar de robustez en aquella red.

Algunos ejemplos de métricas básicas, 7 de 18 (39%) nodos cuya eliminación desconectan la red, esto implica que el grafo posee conectividad $\kappa(G) = 1$. En el caso de las aristas la evaluación no mejora, ya que 11 de 19 (58%) aristas cuya eliminación desconectan la red. Otra métrica de robustez muy usada es edge betweenness centrality. La red tiene un edge betweenness máximo = 81.0 (La arista (Santiago—La Serena)). Esto implica que, si dividimos por la cantidad de pares de nodos existentes en el grafo,

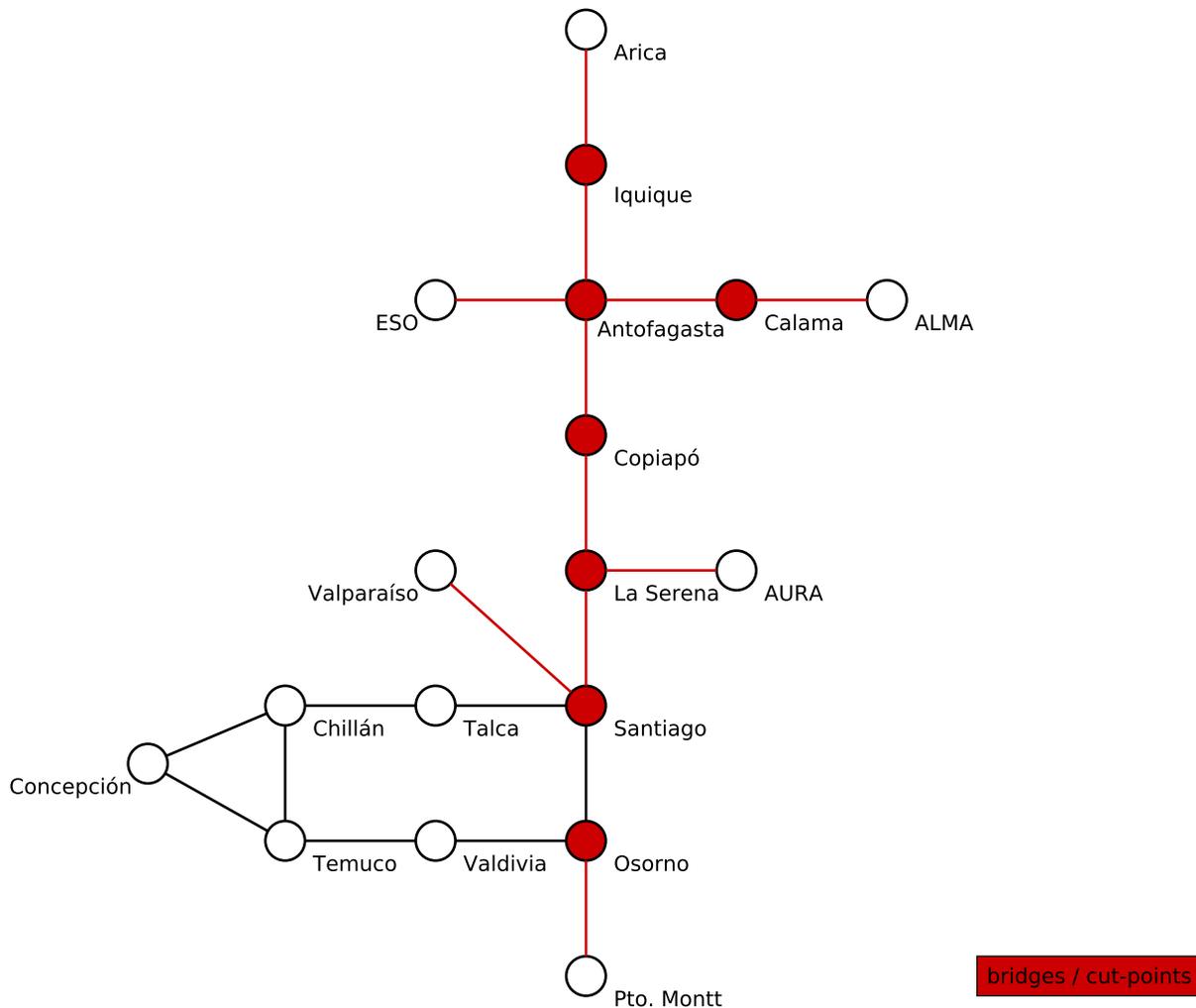


Figura 4.1: Red Universitaria Nacional (REUNA), es una corporación integrada por universidades, centros de investigación y grupos astronómicos internacionales. Los elementos en color rojo, llamados *bridges/cut-points* corresponden a aquellos elementos cuya eliminación desconecta la red.

$$\frac{81.0}{N(N-1)/2} \simeq 0.53, \quad (4.1)$$

significa que 53% de los caminos (o bien más corto, o bien alternativos) se verán afectados por la eliminación de la arista (Santiago—La Serena).

Por último, existen métricas que exigen un **nivel mínimo** de robustez, como el Impacto nodal de Wiener [23].

4.1.1. Métricas elegidas

Dado que la red posee demasiados puntos de falla, se favorece el uso de métricas básicas, que a pesar de ser menos finas, permiten obtener una idea de la robustez de la red. En la siguiente lista se describirán brevemente las métricas usadas.

- *node connectivity* $\kappa(G)$ [38]: el mínimo número de nodos la cual su eliminación desconecta el grafo. Cuando un grafo posee $\kappa(G) = k$, lo llamamos k -conexo.
- *edge betweenness centrality* $c_B(e)$ [35]: propuesta por Freeman (1977). Se calcula como el número de caminos más cortos que pasan a través de tal enlace. Se utiliza la versión en enlaces porque un nodo al ser eliminado, también elimina enlaces lo cual no permite un análisis tan fino.
- *cutset of size k ratio* CS_k [38]: Un conjunto de corte de un grafo conexo se refiere a un conjunto de enlaces que su eliminación resulta un grafo desconectado. Cuando un conjunto de corte tiene k elementos, decimos que tiene un tamaño k . CS_k se define como la proporción de conjuntos de corte de tamaño n , sobre todos los posibles conjuntos de corte de tamaño k , es decir $\binom{n}{k}$.
- *node wiener impact* τ_v [23]: Sea G un grafo 2-conexo. *node wiener impact* de un vértice v se define como:

$$\tau_v = W(G - v) - W(G) + T(v), \quad (4.2)$$

Donde $G - v$ se refiere al grafo obtenido luego de eliminar v de G . $W(G)$ y $T(v)$ son el índice de Wiener y la transmisión del nodo respectivamente, definidas como:

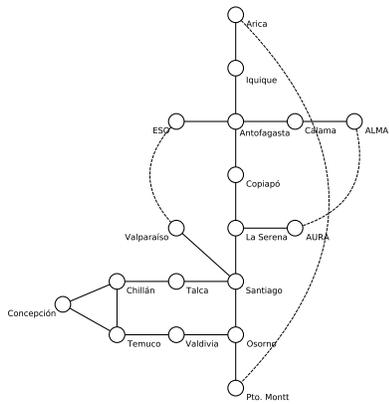
$$T(v) = \sum_{u \in V} D(u, v) \quad (4.3)$$

$$W(G) = \sum_{u \in V} \sum_{v \in V, v < u} D(u, v) \quad (4.4)$$

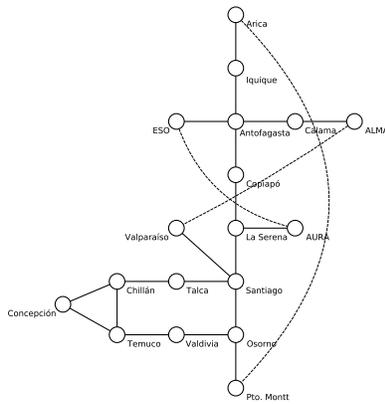
Mide que tanto las distancias totales del grafo son afectadas bajo la eliminación de un nodo.

4.1.2. Primera mejora de la red

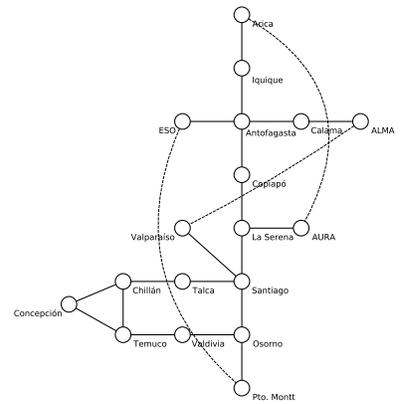
Para conseguir un nivel mínimo de robustez, debemos garantizar que la red siga conectada después de la falla (cualquiera) de algún enlace o nodo, es decir $\kappa(G) > 1$. Por ende, debemos agregar enlaces hasta cumplir con esta propiedad. En REUNA, la cantidad mínima de enlaces requeridos son 3 (conectando aquellos nodos que poseen grado $d(v) = 1$), lo cual nos entregará 6 posibles soluciones, ilustradas, bajo ningún orden en particular, en las figuras 4.2a, 4.2b, 4.2c, 4.2d, 4.2e y 4.2f.



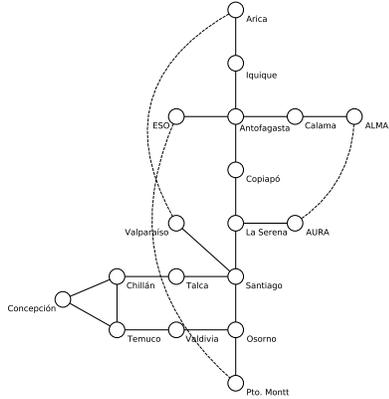
(a) (Valparaíso—ESO), (Arica—Pto. Montt), (AURA—ALMA).



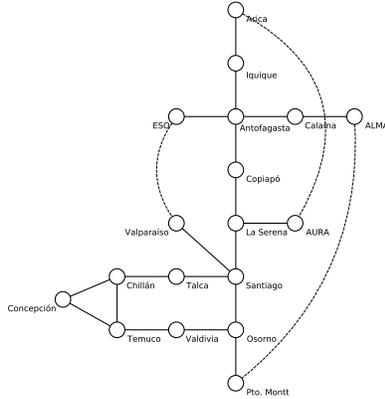
(b) (Valparaíso—ALMA), (Arica—Pto. Montt), (AURA—ESO).



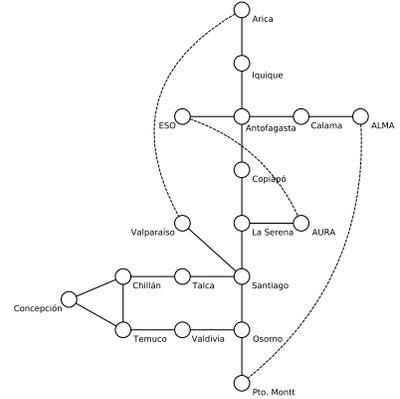
(c) (Valparaíso—ALMA), (Arica—AURA), (Pto. Montt—ESO).



(d) (Valparaíso—Arica), (ALMA—AURA), (Pto. Montt—ESO).



(e) (Valparaíso—ESO), (Arica—AURA), (Pto. Montt—ALMA).



(f) (Valparaíso—Arica), (ESO—AURA), (Pto. Montt—ALMA).

Figura 4.2: Soluciones agregando 3 enlaces.

Estas soluciones, si bien logran que el grafo sea 2-conexo, varían en calidad (en términos de robustez) y mejoran la red. La siguiente tabla comparativa 4.1 muestra el valor de las métricas básicas elegidas.

Tabla 4.1: Comparación de valores de las métricas para las soluciones adicionando 3 enlaces.

| | (a) | (b) | (c) | (d) | (e) | (f) |
|---------|-------|-----|-------|-------|-------|-----|
| CS_2 | 10 | 10 | 10 | 10 | 10 | 10 |
| CS_3 | 31 | 31 | 31 | 31 | 31 | 31 |
| MAX_EBC | 43.89 | 43 | 43.17 | 43.17 | 43.89 | 43 |
| MAX_NWI | 182 | 193 | 161 | 161 | 182 | 193 |

4.1.3. Segunda mejora de la red

Primero comprobamos cual es el mínimo valor posible para esta métrica moviendo cualquiera de los $m + 4$ aristas. Usando el software AutoGraphiX (AGX), buscamos la mejor solución minimizando las métricas definidas. Se propone como heurística minimizar el impacto nodal de Wiener máximo. Esto nos da 2 soluciones posibles (isomorfias ignorando las etiquetas de los nodos) con $\max_v \tau(v) = 41$. Esto nos permite saber qué tan cerca estamos de la posible mejor solución (matemáticamente).

AutoGraphiX (AGX) [11] es un software creado con el objetivo de ayudar a los investigadores en la teoría de grafos. Su principal propósito es la búsqueda de grafos extremales, es decir, grafos que minimizan o maximizan una invariante (o una función de invariantes, que también puede ser considerada como invariante).

AGX utiliza la meta-heurística llamada *Variable Neighbourhood Search* (VNS) [37], usada en muchos campos de la ciencia para resolver problemas de optimización. Este básico algoritmo periódicamente realiza perturbaciones y búsquedas locales de soluciones, manteniendo la mejor solución siempre. Un pseudo código se presenta a continuación:

La sub-rutina de búsqueda local, *LocalSearch*, se refiere a la aplicación sucesiva de cambios, utilizando una *transformación* predeterminada, a la solución presente hasta que no exista mejora posible. Esta solución es llamada óptimo local con respecto a la transformación usada (Por ejemplo, en nuestro caso cambiamos enlaces (aquellos por agregar) de posición para transformar la solución).

Una *transformación*, corresponde a cualquier variación en la solución que, ojalá, cambie la solución lo suficiente para ser analizada. Normalmente es definida como una función de un elemento del conjunto de soluciones, la cual nos entrega un conjunto de soluciones. Por ejemplo, en el caso de los grafos, una transformación puede ser el movimiento, agregación y eliminación de enlaces, que aplicadas sobre un grafo en particular, entreguen un conjunto de grafos con aquellos movimientos de enlaces indicados por la transformación.

En cada iteración, para ahorrar coste computacional y explorar mejores soluciones (y así no quedar atrapados en un óptimo local), perturbamos la mejor solución conocida, *PERTURB*, aplicando la transformación k veces sucesivas.

Algorithm 1 Variable Neighbourhood Search

```
1: procedure VNS( $S, k_{max}$ )
2:    $S^* \leftarrow S$  ▷ Mejor solución conocida
3:    $k \leftarrow 1$ 
4:   while some_criteria do
5:      $S' \leftarrow PERTURB(S^*, k)$ 
6:      $S' \leftarrow LocalSearch(S')$ 
7:     if  $S'$  better than  $S^*$  then
8:        $S^* \leftarrow S'$ 
9:        $k \leftarrow 1$ 
10:    else
11:       $k \leftarrow k + 1$ 
12:      if  $k > k_{max}$  then
13:         $k \leftarrow 1$ 
14:    return  $S^*$ 
```

Desafortunadamente, AGX no se pueden construir grafos de orden mayor a 25 nodos, por ende, para usos futuros de la heurística, se implementó un software propio disponible online¹

Para encontrar la mejor solución, nos interesa minimizar cada una de las métricas, es decir, el impacto nodal de Wiener máximo (MAX_NWI), el edge betweenness centrality máximo (MAX_EBC), y la razón de la cantidad de conjuntos de corte de tamaño dos versus la cantidad total de conjuntos de aristas de tamaño 2 (Se puede considerar conjuntos de cortes más grandes, pero dado el coste computacional asociado $O(k^m)$ se eligió el mínimo $k = 2$). Sólo está permitido mover los 4 enlaces adicionales. Las soluciones junto con la mínima posible se presentan en la figura 4.3

Como el grafo REUNA es lo suficientemente pequeño, podemos observar todas las soluciones posibles y estudiar su correlación. La figura 4.4 presenta las soluciones ordenadas por el impacto de Wiener, dado que *edge betweenness* y la cantidad de conjuntos de corte crece a lo largo del gráfico se puede concluir que existe una leve correlación entre aquellas métricas, lo cual implica que minimizar el impacto nodal de Wiener tiene un efecto positivo con respecto a las otra métricas, dada la topología de estudio. La figura 4.5 presente lo mismo pero en escala logarítmica.

4.1.4. Discusión

La tabla 4.2 nos muestra comparativamente que tanto se mejora la solución agregando un enlace adicional. Con el coste adicional de agregar un enlace podemos ganar una mejora significativa en cada una de las métricas (ver figura 4.3). En el caso del impacto nodal de Wiener, podemos observar que logra reducir la métrica a un 39% de la mejor solución anterior (c) (ver 4.2c) o (d) (ver 4.2d).

En la figura 4.3 podemos observar cuál es la mejora entre una solución y otra. La figura

¹https://github.com/niclabs/resilience_metrics

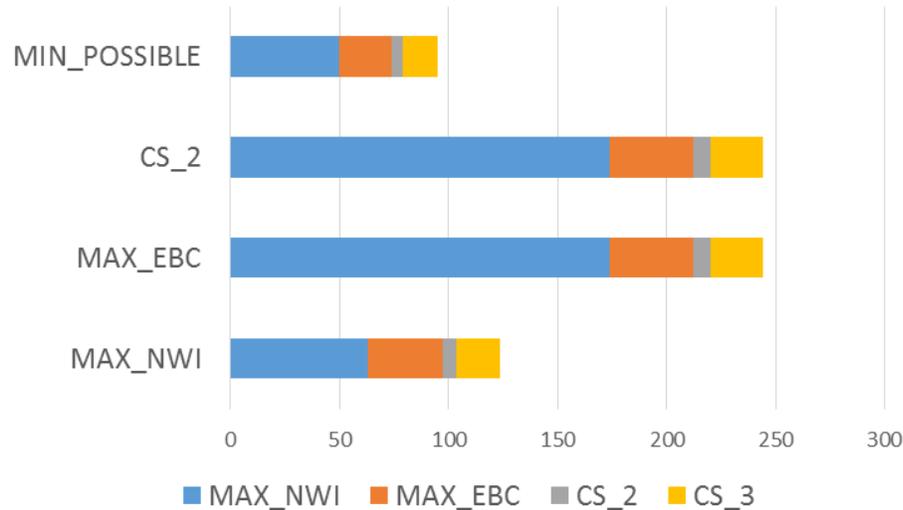


Figura 4.3: El rendimiento de cada solución según cada métrica seleccionada a minimizar con VNS. (MIN POSSIBLE) se refiere a los mínimos globales de cada métrica por separado (encontradas por “fuerza bruta”, no es aplicable a grafos más grande). Notar que una solución que tenga el mínimo global para una métrica no necesariamente minimiza globalmente otras.

4.6 muestra la solución encontrada, cercana a la solución (e) (ver 4.2e), la cual es una de las peores soluciones agregando 3 enlaces. Minimizando MAX_NWI se desempeña mejor ya que logra obtener mejores valores en todas las otras métricas. Aunque esta solución requiere de enlaces geográficamente largos (vale decir que existe un cable submarino entre “Arica” y “Valparaíso”). Notar que las soluciones minimizando las otras dos métricas son óptimos locales.

Si bien esta solución no necesariamente es el óptimo global con respecto a alguna métrica, las soluciones indican que minimizar el impacto de Wiener tiene un efecto positivo sobre las otras métricas estudiadas. Cabe mencionar que la nueva solución no corresponde al óptimo global en función del impacto de Wiener.

Tabla 4.2: Valores de métricas de la solución escogida.

| | (a) | (b) | (c) | (d) | (e) | (f) | Nueva solución |
|---------|-------|-----|-------|-------|-------|-----|-----------------------|
| CS_2 | 10 | 10 | 10 | 10 | 10 | 10 | 6 |
| CS_3 | 31 | 31 | 31 | 31 | 31 | 31 | 20 |
| MAX_EBC | 43.89 | 43 | 43.17 | 43.17 | 43.89 | 43 | 34.33 |
| MAX_NWI | 182 | 193 | 161 | 161 | 182 | 193 | 63 |

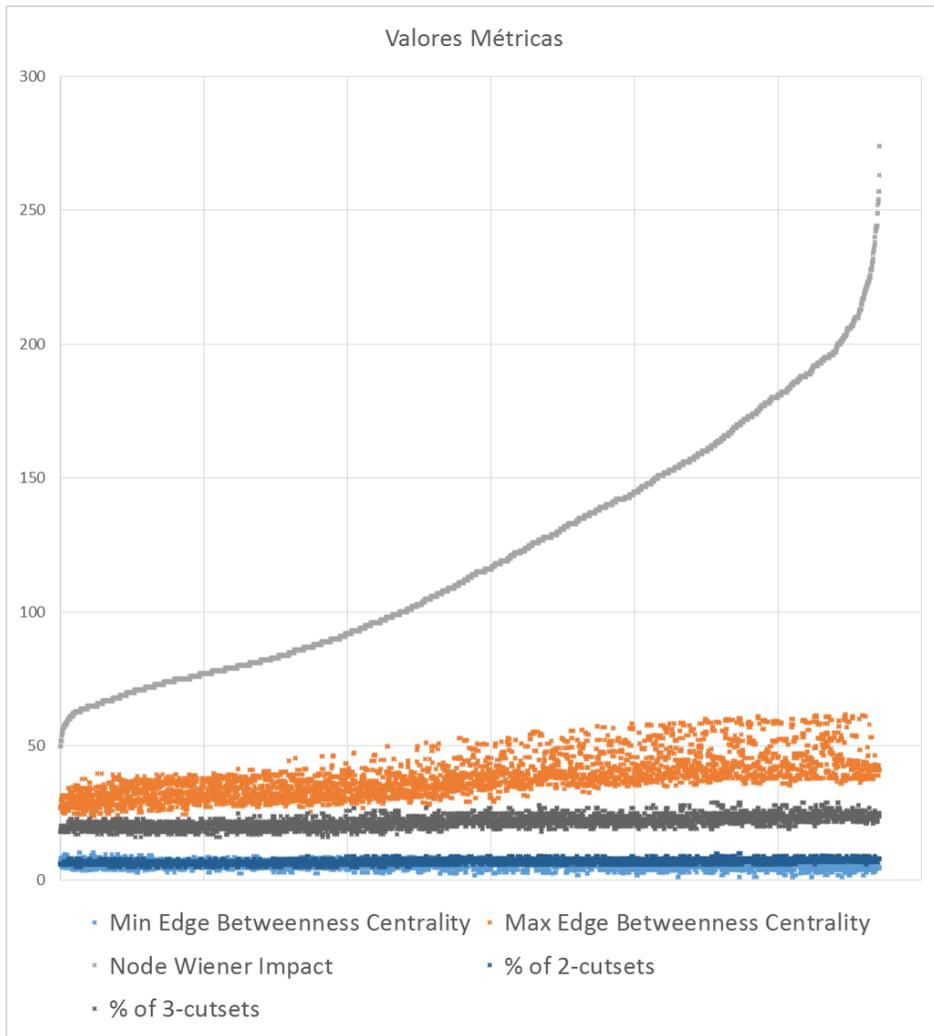


Figura 4.4: Valores de las métricas por cada solución, ordenadas por el impacto nodal de Wiener.

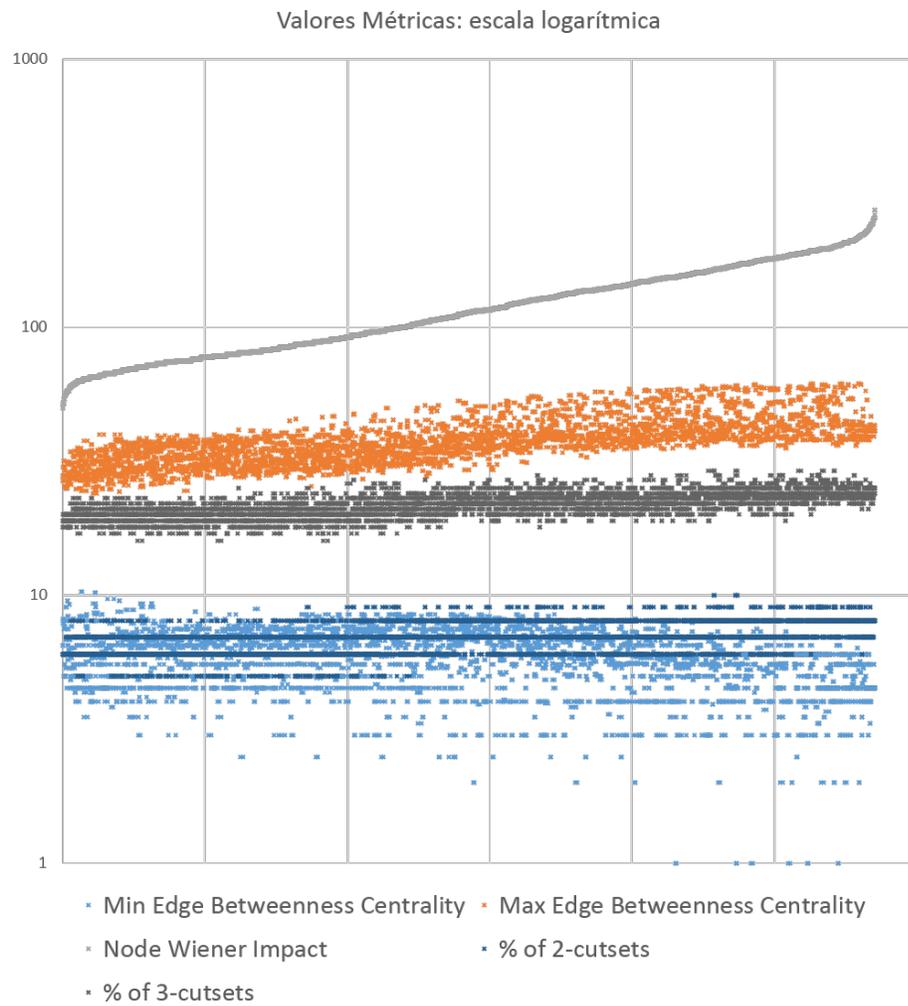


Figura 4.5: Valores de las métricas por cada solución, ordenadas por el impacto nodal de Wiener. En escala logarítmica.

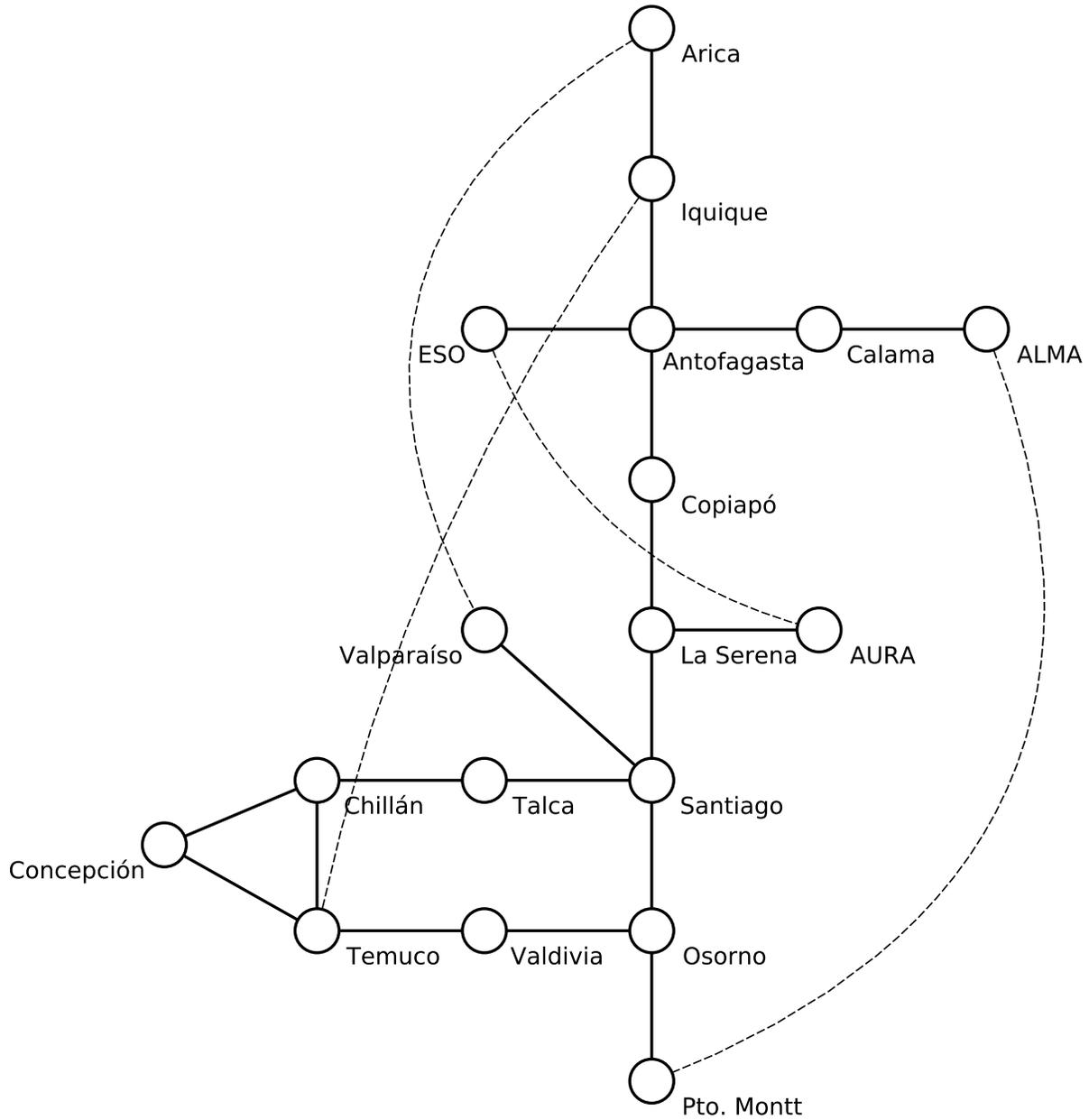


Figura 4.6: Solución encontrada minimizando MAX_NWI , con 4 enlaces adicionales.

Capítulo 5

Conclusiones y trabajo futuro

5.1. Conclusiones

Los resultados del mapeo sistemático indican una auge en el desarrollo de las métricas en los últimos años, y lamentablemente la mayoría de las propuestas tienden a ser complejas y poco aplicables a grafos simples e intrínsecamente poco robustos. No se encontraron métricas que, por naturaleza, fuesen aplicables a modelos simples y frágiles (aquellos con conectividad igual a 1), ya que, probablemente, este caso no se ha considerado lo suficientemente interesante para aplicaciones reales.

Así también, en consecuencia, observando la topología posible y las restricciones de territorio, la mayoría de las métricas dan resultados poco favorables para el ejemplo planteado (REUNA) y se conjetura, con fuertes bases, que tampoco serán favorables para el grafo real de la red física chilena, debido a que ésta también presenta una topología similar (árboles con raíces conectadas con un poco de redundancia).

Por otro lado, si existen métricas, un poco más básicas y antiguas, que logran capturar de algún modo estas vulnerabilidades. Un clásico ejemplo es *betweenness centrality* que logra capturar la importancia de un nodo (o arista) en términos de los caminos que pasan por él. También, tenemos la proporción de conjuntos de corte que desconectan el grafo, que describen claramente la conectividad del grafo en términos de sus enlaces (notar que una falla de un nodo se puede simular como la falla de sus enlaces). Lo negativo es que la métrica tiene un coste exponencial lo que desalienta su uso en grafos más grandes. Además, tenemos la centralidad de enlaces, que indica que tan importante son los enlaces en términos de la cantidad de caminos más cortos, pero no necesariamente indica robustez en términos de conectividad. Ya que no toma en cuenta, ni la existencia ni el largo de, los caminos secundarios que puedan ser usar. Afortunadamente, si existe una métrica, el impacto nodal de Wiener, que toma en cuenta los largos de los caminos (y su diferencia con el camino alternativo) y se puede usar para evaluar la robustez del grafo, suponiendo que es al menos 2-conexo.

Los resultados indican, para la topología de nuestro caso de estudio, que existe una correlación del impacto nodal de Wiener con las otras dos métricas, pero sobre todo con la

centralidad de enlaces. Lo cual significa que minimizar el impacto en función de sus enlaces produce resultados favorables para el grafo en cuestión.

Ahora, con respecto a las mejoras, existe una mejora mínima como aquella que vuelva el grafo resistente a cualquier falla aleatoria. Suponiendo fuertemente que el coste de cada enlace es el mismo, agregar la cantidad mínima de enlaces para convertir el grafo a uno 2-conexo se considera como requisito mínimo para hablar de robustez. Luego, la segunda mejora expande el universo de soluciones a uno mucho más grande, que a coste de un enlace adicional, puede mejorar la robustez red en un porcentaje significativo, que además, si eliminamos el supuesto del coste de las aristas, puede ser menos costosa porque las aristas elegidas tienen mayor holgura para cumplir con el requisito mínimo de robustez. En este sentido, las mejoras nos permiten determinar qué métrica es mejor utilizar a la hora de usar la red real, usando la metodología de la segunda mejora.

Con respecto a la pregunta de investigación inicial, este trabajo de tesis logra reconocer y resumir la literatura científica a través de la última década, correspondiente al tema de las métricas de robustez en redes complejas. Adicionalmente, dado que el impacto de Wiener presenta correlaciones favorables con las otras métricas evaluadas, se concluye que esta métrica es útil para medir robustez en redes con escasa o nula redundancia. Por ende, esta métrica también amerita estudio adicional.

5.2. Trabajo futuro

Como trabajo futuro, se puede determinar más correlaciones y efectos sobre métricas adicionales. Un ejemplo de métrica, es *betweenness centrality*, una métrica ampliamente usada. En general, un bajo valor máximo indica buena robustez y se puede estudiar el efecto del impacto nodal de Wiener sobre esta métrica y viceversa.

Además, también corresponde aplicar la mejora a la red real y estudiar sus efectos. La cantidad de enlaces necesaria para convertir la red a una 2-conexa será mucho mayor y por ende la cantidad de soluciones iniciales será mayor. Claramente, la cantidad de soluciones de la segunda mejora será extremadamente mayor y estudiar cada una de las soluciones no será posible, pero estudiar el rendimiento de la heurística y sus efectos también es una posible dirección.

Finalmente, sin menor importancia, extender estas soluciones a aristas con costes. Estos costes probablemente serán proporcionales a las distancias geográficas e influidas por asuntos similares, como la conservación de zonas naturales, características del terreno, entre otros.

Capítulo 6

Bibliografía

- [1] Réka Albert, Hawoong Jeong, and Albert-László Barabási. Error and attack tolerance of complex networks. *Nature*, 406:378–382, July 2000.
- [2] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.
- [3] Albert-László Barabási, Réka Albert, and Hawoong Jeong. Mean-field theory for scale-free random networks. *Physica A*, 272:173–187, 1999.
- [4] Ashkan Behnam, Gijs Bosman, and Ant Ural. Percolation scaling of 1/f noise in single-walled carbon nanotube films. *Physical Review B*, 78(8):085431, 2008.
- [5] Alina Beygelzimer, Geoffrey Grinstein, Ralph Linsker, and Irina Rish. Improving network robustness by edge modification. *Physica A: Statistical Mechanics and its Applications*, 357:593–612, November 2005.
- [6] Ramesh Bhandari. *Survivable Networks: Algorithms for Diverse Routing*. Kluwer Academic Publishers, Norwell, MA, USA, 1998.
- [7] Kashif Bilal, Marc Manzano, Samee U Khan, Eusebi Calle, Keqin Li, and Albert Y Zomaya. On the characterization of the structural robustness of data center networks. *IEEE Transactions on Cloud Computing*, 1(1):1–1, 2013.
- [8] Béla Bollobás and Oliver Riordan. Robustness and vulnerability of scale-free random graphs. *Internet Mathematics*, 1:1–35, 2003.
- [9] Sergey Brin and Lawrence Page. The anatomy of a large-scale hypertextual web search engine. *Computer networks and ISDN systems*, 30(1):107–117, 1998.
- [10] Duncan S. Callaway, M. E. J. Newman, Steven H. Strogatz, and Duncan J. Watts. Network robustness and fragility: Percolation on random graphs. *Physical Review Letters*, 85:5468, December 2000.
- [11] Gilles Caporossi and Pierre Hansen. Variable neighborhood search for extremal graphs:

- 1 the autographix system. *Discrete Mathematics*, 212(1-2):29–44, 2000.
- [12] Yiping Chen, Gerald Paul, Reuven Cohen, Shlomo Havlin, Stephen P Borgatti, Fredrik Liljeros, and H Eugene Stanley. Percolation theory applied to measures of fragmentation in social networks. *Physical Review E*, 75(4):046107, 2007.
- [13] Xue-Qi Cheng, Fu-Xin Ren, Hua-Wei Shen, Zi-Ke Zhang, and Tao Zhou. Bridgeness: a local index on edge significance in maintaining global connectivity. *Journal of Statistical Mechanics: Theory and Experiment*, 2010(10):P10011, 2010.
- [14] Yufei Cheng, M Todd Gardner, Junyan Li, Rebecca May, Deep Medhi, and James PG Sterbenz. Analysing geopath diversity and improving routing performance in optical networks. *Computer Networks*, 82:50–67, 2015.
- [15] Piotr Cholda, Anders Mykkeltveit, Bjarne E. Helvik, Otto J. Wittner, and Andrzej Jajszczyk. A survey of resilience differentiation frameworks in communication networks. *IEEE Communications Surveys & Tutorials*, 9:32–55, October 2007.
- [16] Fan R. K. Chung. *Spectral Graph Theory*, volume 92 of *CBMS Regional Conference Series On Mathematics*. American Mathematical Society, 1997.
- [17] Reuven Cohen, Keren Erez, Daniel ben Avraham, and Shlomo Havlin. Resilience of the internet to random breakdowns. *Physical Review Letters*, 85:4626, November 2000.
- [18] Reuven Cohen, Keren Erez, Daniel ben Avraham, and Shlomo Havlin. Breakdown of the internet under intentional attack. *Physical Review Letters*, 86:3682, April 2001.
- [19] Carlos Correa, Tarik Crnovrsanin, and Kwan-Liu Ma. Visual reasoning about social networks using centrality sensitivity. *IEEE transactions on visualization and computer graphics*, 18(1):106–120, 2012.
- [20] Anthony H Dekker and Bernard D Colbert. Network robustness and graph topology. In *Proceedings of the 27th Australasian conference on Computer science-Volume 26*, pages 359–368. Australian Computer Society, Inc., 2004.
- [21] Anthony H. Dekker and Bernard D. Colbert. Network robustness and graph topology. In *Proc. of the 27th Australasian conference on Computer science*, volume 26, pages 359–368, Darlinghurst, Australia, 2004. Australian Computer Society, Inc.
- [22] Jean-Charles Delvenne and Anne-Sophie Libert. Centrality measures and thermodynamic formalism for complex networks. *Physical Review E*, 83(4):046117, 2011.
- [23] C.K. Dominicini, M.H.M Paiva, G. Caporossi, M.A. Marotta, M.R.N. Ribeiro, N.J. Kaminski, M.E.V. Segatto, M. Martinello, C.B. Both, J.M. Marquez-Barja, and L.A. DaSilva. Optimizing C-RAN backhaul topologies: A resilience-oriented approach using graph invariants. *Les Cahiers du GERAD*, G-2016-57:1–16, 2016.
- [24] Jun Dong and Steve Horvath. Understanding network concepts in modules. *BMC systems biology*, 1(1):24, 2007.

- [25] Arnold Dresden. The fourteenth western meeting of the american mathematical society. *Bulletin of the American Mathematical Society*, 26(9):385–396, 1920.
- [26] David Easley and Jon Kleinberg. *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge University Press, 2010.
- [27] Pál Erdős and Alfréd Rényi. On random graphs I. In *Selected papers of Alfréd Rényi*, volume 2, pages 308–315. Akadémiai Kiadó, Budapest, Hungary, 1976. First publication in 1959.
- [28] Pál Erdős and Alfréd Rényi. On the evolution of random graphs. In *Selected papers of Alfréd Rényi*, volume 2, pages 482–525. Akadémiai Kiadó, Budapest, Hungary, 1976. First publication in 1960.
- [29] E. Estrada. Network robustness to targeted attacks. the interplay of expansibility and degree distribution. *The European Physical Journal B*, 52:563–574, August 2006.
- [30] Ernesto Estrada. Characterization of 3d molecular structure. *Chemical Physics Letters*, 319(5):713–718, 2000.
- [31] Ernesto Estrada. Characterization of 3d molecular structure. *Chemical Physics Letters*, 319(5):713–718, 2000.
- [32] Ernesto Estrada. Characterization of the folding degree of proteins. *Bioinformatics*, 18(5):697–704, 2002.
- [33] Fenling Feng and Lei Wang. Robustness measure of china’s railway network topology using relative entropy. *Discrete Dynamics in Nature and Society*, 2013, 2013.
- [34] Miroslav Fiedler. Algebraic connectivity of graphs. *Czechoslovak mathematical journal*, 23(2):298–305, 1973.
- [35] Linton C Freeman. A set of measures of centrality based on betweenness. *Sociometry*, pages 35–41, 1977.
- [36] E. N. Gilbert. Random graphs. *Annals of Mathematical Statistics*, 30(4):1141–1144, 1959.
- [37] Pierre Hansen, Nenad Mladenovic, and Jose Andrés Moreno Pérez. Variable neighbourhood search. *Inteligencia Artificial, Revista Iberoamericana de Inteligencia Artificial*, 19:77–92, 2003.
- [38] Frank Harary. *Graph theory*. Addison-Wesley Publishing Company, Philippines, 1969.
- [39] Hong Huang and John A. Copeland. A series of Hamiltonian cycle-based solutions to provide simple and scalable mesh optical network resilience. *IEEE Communications Magazine*, 40:46–51, November 2002.
- [40] Staffs Keele. Guidelines for performing systematic literature reviews in software engi-

neering. In *Technical report, Ver. 2.3 EBSE Technical Report. EBSE*. 2007.

- [41] Jon M Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM (JACM)*, 46(5):604–632, 1999.
- [42] Yakup Koç, Martijn Warnier, Robert E Kooij, and Frances MT Brazier. An entropy-based metric to quantify the robustness of power grids against cascading failures. *Safety science*, 59:126–134, 2013.
- [43] Yakup Koç, Martijn Warnier, Piet Van Mieghem, Robert E Kooij, and Frances MT Brazier. The impact of the topology on cascading failures in a power grid model. *Physica A: Statistical Mechanics and its Applications*, 402:169–179, 2014.
- [44] Vito Latora and Massimo Marchiori. Efficient behavior of small-world networks. *Physical Review Letters*, 87(19):198701, November 2001.
- [45] Li Li, Qing-Shan Jia, Hengtao Wang, Ruixi Yuan, and Xiaohong Guan. A systematic method for network topology reconfiguration with limited link additions. *Journal of Network and Computer Applications*, 35(6):1979–1989, 2012.
- [46] Dajie Liu, Huijuan Wang, and Piet Van Mieghem. Spectral perturbation and reconstructability of complex networks. *Physical Review E*, 81(1):016101, 2010.
- [47] Lijia Ma, Maoguo Gong, Qing Cai, and Licheng Jiao. Enhancing community integrity of networks against multilevel targeted attacks. *Physical Review E*, 88(2):022810, 2013.
- [48] Priya Mahadevan, Dmitri Krioukov, Marina Fomenkov, Xenofontas Dimitropoulos, Amin Vahdat, et al. The internet as-level topology: three data sources and one definitive metric. *ACM SIGCOMM Computer Communication Review*, 36(1):17–26, 2006.
- [49] Priya Mahadevan, Dmitri Krioukov, Marina Fomenkov, Bradley Huffaker, Xenofontas Dimitropoulos, K. C. Claffy, and Amin Vahdat. The internet AS-level topology: three data sources and one definitive metric. *ACM SIGCOMM Computer Communication Review*, 36:17–26, January 2006.
- [50] Fragkiskos D Malliaros, Vasileios Megalooikonomou, and Christos Faloutsos. Estimating robustness in large social graphs. *Knowledge and Information Systems*, 45(3):645–678, 2015.
- [51] Alberto Medina, Ibrahim Matta, and John Byers. On the origin of power laws in Internet topologies. *ACM Computer Communication Review*, 30(2):18–28, 2000.
- [52] Edward K. Morlok and David J. Chang. Measuring capacity flexibility of a transportation system. *Transportation Research Part A: Policy and Practice*, 38:405–420, July 2004.
- [53] Pamela M. Murray-Tuite. A comparison of transportation network resilience under simulated system optimum and user equilibrium conditions. In L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, and D. M. Nicoland R. M. Fujimoto, editors, *Proc. of*

the 2006 Winter Simulation Conference, pages 1398–1405. IEEE, 2006.

- [54] Mark E. J. Newman. A measure of betweenness centrality based on random walks. Technical Report cond-mat/0309045, arXiv.org e-Print archive, <http://arxiv.org/>, 2003.
- [55] Mark E. J. Newman and Michelle Girvan. Mixing patterns and community structure in networks. In Romualdo Pastor-Satorras, Miguel Rubi, and Albert Diaz-Guilera, editors, *Statistical Mechanics of Complex Networks*, volume 625 of *Lecture Notes in Physics*, pages 66–87, Berlin, Germany, 2003. Springer-Verlag GmbH.
- [56] Mark E. J. Newman and Duncan J. Watts. Scaling and percolation in the small-world network model. *Physical Review E*, 60(6):7332–7342, 1999.
- [57] David C Novak, James L Sullivan, and Darren M Scott. A network-based approach for evaluating and ranking transportation roadway projects. *Applied geography*, 34:498–506, 2012.
- [58] Kuang-Chang Pien, Ke Han, Wenlong Shang, Arnab Majumdar, and Washington Ochieng. Robustness analysis of the european air traffic network. *Transportmetrica A: Transport Science*, 11(9):772–792, 2015.
- [59] Mahendra Piraveenan, Mikhail Prokopenko, and Liaquat Hossain. Percolation centrality: Quantifying graph-theoretic impact of nodes during percolation in networks. *PloS one*, 8(1):e53095, 2013.
- [60] Éva E Plagányi, Ingrid Van Putten, Olivier Thébaud, Alistair J Hobday, James Innes, Lilly Lim-Camacho, Ana Norman-López, Rodrigo H Bustamante, Anna Farmery, Aysha Fleming, et al. A quantitative metric to identify critical elements within seafood supply networks. *PloS one*, 9(3):e91833, 2014.
- [61] Anatolij P Potapov, Björn Goemann, and Edgar Wingender. The pairwise disconnectivity index as a new metric for the topological analysis of regulatory networks. *BMC bioinformatics*, 9(1):227, 2008.
- [62] AP Quayle, AS Siddiqui, and SJM Jones. Preferential network perturbation. *Physica A: Statistical Mechanics and its Applications*, 371(2):823–840, 2006.
- [63] Daniel J Rosenkrantz, Sanjay Goel, SS Ravi, and Jagdish Gangolly. Resilience metrics for service-oriented networks: a service allocation approach. *IEEE Transactions on Services Computing*, 2(3):183–196, 2009.
- [64] Eduardo Sáenz-de Cabezón and Henry P Wynn. Measuring the robustness of a network using minimal vertex covers. *Mathematics and Computers in Simulation*, 104:82–94, 2014.
- [65] Ronaldo M Salles and Donato A Marino. Strategies and metric for resilience in computer networks. *The Computer Journal*, page bxr110, 2011.

- [66] Giovanni Scardoni and Carlo Laudanna. Identifying critical traffic jam areas with node centralities interference and robustness. *NETWORKS AND HETEROGENEOUS MEDIA*, 7(3):463–471, 2012.
- [67] Salvatore Scellato, Ilias Leontiadis, Cecilia Mascolo, Prithwish Basu, and Murtaza Zaffer. Evaluating temporal robustness of mobile networks. *IEEE Transactions on Mobile Computing*, 12(1):105–117, 2013.
- [68] Alberto Schaeffer-Filho, Paul Smith, Andreas Mauthe, David Hutchison, Yue Yu, and Michael Fry. A framework for the design and evaluation of network resilience management. In *Proc. of the IEEE Network Operations and Management Symposium*, pages 401–408. IEEE, 2012.
- [69] Christian M Schneider, André A Moreira, José S Andrade, Shlomo Havlin, and Hans J Herrmann. Mitigation of malicious attacks on networks. *Proceedings of the National Academy of Sciences*, 108(10):3838–3841, 2011.
- [70] Alexander Schrijver. On the history of the transportation and maximum flow problems. *Mathematical Programming*, 91(3):437–445, 2002.
- [71] Darren M. Scott, David C. Novak, Lisa Aultman-Hall, and Feng Guo. Network robustness index: A new method for identifying critical links and evaluating the performance of transportation networks. *Journal of Transport Geography*, 14:215–227, May 2006.
- [72] Paul Smith, David Hutchison, James P. G. Sterbenz, Marcus Schöller, Ali Fessi, Merkuris Karaliopoulos, Chidung Lac, and Bernhard Plattner. Network resilience: A systematic approach. *IEEE Communications Magazine*, 49:88–97, July 2011.
- [73] James P.G. Sterbenz, Egemen K. Çetinkaya, Mahmood A. Hameed, Abdul Jabbar, Shi Qian, and Justin P. Rohrer. Evaluation of network resilience, survivability, and disruption tolerance: analysis, topology generation, simulation, and experimentation. *Telecommunication Systems*, 52:705–736, February 2013.
- [74] Ali Sydney, Caterina Scoglio, Phillip Schumm, and Robert E Kooij. Elasticity: topological characterization of robustness in complex networks. In *Proceedings of the 3rd International Conference on Bio-Inspired Models of Network, Information and Computing Systems*, page 19. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2008.
- [75] Somnath Tagore and Rajat K De. Detecting breakdown points in metabolic networks. *Computational biology and chemistry*, 35(6):371–380, 2011.
- [76] Xianglong Tang, Jing Liu, and Mingxing Zhou. Enhancing network robustness against targeted and random attacks using a memetic algorithm. *EPL (Europhysics Letters)*, 111(3):38005, 2015.
- [77] T. Tanizawa, G. Paul, R. Cohen, S. Havlin, and H. E. Stanley. Optimization of network robustness to waves of targeted and random attacks. *Physical Review E*, 71:047101, 2005.

- [78] Ali Tizghadam and Alberto Leon-Garcia. Autonomic traffic engineering for network robustness. *IEEE journal on selected areas in communications*, 28(1):39–50, 2010.
- [79] Ali Tizghadam and Alberto Leon-Garcia. Autonomic traffic engineering for network robustness. *IEEE Journal on Selected Areas in Communications*, 28:39–50, January 2010.
- [80] Rostislav Vodák, Michal Bíl, and Jiří Sedoník. Network robustness and random processes. *Physica A: Statistical Mechanics and its Applications*, 428:368–382, 2015.
- [81] Bing Wang, Huanwen Tang, Chonghui Guo, and Zhilong Xiu. Entropy optimization of scale-free networks’ robustness to random failures. *Physica A: Statistical Mechanics and its Applications*, 363(2):591–596, 2006.
- [82] Jian-Wei Wang and Li-Li Rong. A model for cascading failures in scale-free networks with a breakdown probability. *Physica A: Statistical Mechanics and its Applications*, 388(7):1289–1298, 2009.
- [83] John Glen Wardrop. Road paper. some theoretical aspects of road traffic research. In *ICE Proceedings: engineering divisions*, volume 1, pages 325–362. Thomas Telford, 1952.
- [84] Duncan J. Watts and Steven H. Strogatz. Collective dynamics of ’small world’ networks. *Nature*, 393:440–442, 1998.
- [85] Jun Wu, Mauricio Barahona, Yue-Jin Tan, and Hong-Zhong Deng. Spectral measure of structural robustness in complex networks. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 41(6):1244–1252, 2011.
- [86] Yan Xin and Wu Yang. Topological and spectral perturbations in complex networks. *Chinese Physics Letters*, 29(12):128901, 2012.
- [87] Ting Yang, Zhixian Lin, and Bo Yuan. A betweenness calibration topology optimal control algorithm for wireless sensor networks. *International Journal of Distributed Sensor Networks*, 2013, 2013.
- [88] Shang Yi-Lun. Local natural connectivity in complex networks. *Chinese Physics Letters*, 28(6):068903, 2011.
- [89] Mina Youssef, Robert Kooij, and Caterina Scoglio. Viral conductance: Quantifying the robustness of networks with respect to spread of epidemics. *Journal of Computational Science*, 2(3):286–298, 2011.
- [90] An Zeng and Weiping Liu. Enhancing network robustness against malicious attacks. *Physical Review E*, 85(6):066130, 2012.
- [91] Jianhua Zhang, Xiaoming Xu, Liu Hong, Shuliang Wang, and Qi Fei. Networked analysis of the shanghai subway network, in china. *Physica A: Statistical Mechanics and its Applications*, 390(23):4562–4570, 2011.

- [92] Kang Zhao, Akhil Kumar, Terry P. Harrison, and John Yen. Analyzing the resilience of complex supply network topologies against random and targeted disruptions. *IEEE Systems Journal*, 5:28–39, March 2011.
- [93] Kang Zhao, Akhil Kumar, and John Yen. Achieving high robustness in supply distribution networks by rewiring. *IEEE Transactions on Engineering Management*, 58(2):347–362, 2011.

Capítulo 7

Anexos

7.1. Consultas realizadas

- ACM Digital Library.

```
recordAbstract:(+"network" +"robustness metric")
recordAbstract:(+"network" +"robustness indicator")
recordAbstract:(+"network" +"robustness measure")
recordAbstract:(+"network" +"robustness index")
recordAbstract:(+"network" +"percolation metric")
recordAbstract:(+"network" +"percolation indicator")
recordAbstract:(+"network" +"percolation measure")
recordAbstract:(+"network" +"percolation index")
recordAbstract:(+"network" +"resilience metric")
recordAbstract:(+"network" +"resilience indicator")
recordAbstract:(+"network" +"resilience measure")
recordAbstract:(+"network" +"resilience index")
acmdlTitle:(+"network" +"robustness metric")
acmdlTitle:(+"network" +"robustness indicator")
acmdlTitle:(+"network" +"robustness measure")
acmdlTitle:(+"network" +"robustness index")
acmdlTitle:(+"network" +"percolation metric")
acmdlTitle:(+"network" +"percolation indicator")
acmdlTitle:(+"network" +"percolation measure")
acmdlTitle:(+"network" +"percolation index")
acmdlTitle:(+"network" +"resilience metric")
acmdlTitle:(+"network" +"resilience indicator")
acmdlTitle:(+"network" +"resilience measure")
acmdlTitle:(+"network" +"resilience index")
```

- IEEE Xplore (IEEE Electronic Library - IEL).

La consulta debe ser dividida ya que el motor de búsqueda de IEE no acepta más de 15 términos.

"Abstract":(((robustness ONEAR/0 metric) OR
(robustness ONEAR/0 measure) OR (robustness ONEAR/0 indicator) OR
(robustness ONEAR/0 index)) AND network)

"Document Title":(((robustness ONEAR/0 metric) OR
(robustness ONEAR/0 measure) OR (robustness ONEAR/0 indicator) OR
(robustness ONEAR/0 index)) AND network)

"Abstract":(((robustness ONEAR/0 metrics) OR
(robustness ONEAR/0 measures) OR (robustness ONEAR/0 indicators) OR
(robustness ONEAR/0 indexes)) AND network)

"Document Title":(((robustness ONEAR/0 metrics) OR
(robustness ONEAR/0 measures) OR (robustness ONEAR/0 indicators) OR
(robustness ONEAR/0 indexes)) AND network)

"Abstract":(((percolation ONEAR/0 metric) OR
(percolation ONEAR/0 measure) OR (percolation ONEAR/0 indicator) OR
(percolation ONEAR/0 index)) AND network)

"Document Title":(((percolation ONEAR/0 metric) OR
(percolation ONEAR/0 measure) OR (percolation ONEAR/0 indicator) OR
(percolation ONEAR/0 index)) AND network)

"Abstract":(((percolation ONEAR/0 metrics) OR
(percolation ONEAR/0 measures) OR (percolation ONEAR/0 indicators) OR
(percolation ONEAR/0 indexes)) AND network)

"Document Title":(((percolation ONEAR/0 metrics) OR
(percolation ONEAR/0 measures) OR (percolation ONEAR/0 indicators) OR
(percolation ONEAR/0 indexes)) AND network)

"Abstract":(((resilience ONEAR/0 metric) OR
(resilience ONEAR/0 measure) OR (resilience ONEAR/0 indicator) OR
(resilience ONEAR/0 index)) AND network)

"Document Title":(((resilience ONEAR/0 metric) OR
(resilience ONEAR/0 measure) OR (resilience ONEAR/0 indicator) OR
(resilience ONEAR/0 index)) AND network)

"Abstract":(((resilience ONEAR/0 metrics) OR
(resilience ONEAR/0 measures) OR (resilience ONEAR/0 indicators) OR
(resilience ONEAR/0 indexes)) AND network)

"Document Title":(((resilience ONEAR/0 metrics) OR
(resilience ONEAR/0 measures) OR (resilience ONEAR/0 indicators) OR
(resilience ONEAR/0 indexes)) AND network)

- ScienceDirect.


```
tak({percolation metric} OR {resilience metric} OR
{robustness metric}) AND tak({network})

tak({percolation measure} OR {resilience measure} OR
{robustness measure}) AND tak({network})

tak({percolation indicator} OR {resilience indicator} OR
{robustness indicator}) AND tak({network})

tak({percolation index} OR {resilience index} OR
{robustness index}) AND tak({network})
```
- Web of Science.


```
TS=(complex NEAR network) AND
TS=((percolation OR resilience OR robust) NEAR
(metric OR measure OR indicator OR index))
```
- PLoS (Public Library of Science).


```
(abstract:"network") AND
(abstract:"robustness measure" OR abstract:"robustness metric" OR
abstract:"robustness indicator" OR abstract:"robustness index" OR
abstract:"resilience measure" OR abstract:"resilience metric" OR
abstract:"resilience indicator" OR abstract:"resilience index" OR
abstract:"percolation measure" OR abstract:"percolation metric" OR
abstract:"percolation indicator" OR abstract:"percolation index")

(title:"network") AND
(title:"robustness measure" OR title:"robustness metric" OR
title:"robustness indicator" OR title:"robustness index" OR
title:"resilience measure" OR title:"resilience metric" OR
title:"resilience indicator" OR title:"resilience index" OR
title:"percolation measure" OR title:"percolation metric" OR
title:"percolation indicator" OR title:"percolation index")
```
- Nature


```
network "robustness metric"
network "robustness indicator"
network "robustness measure"
network "robustness index"
network "percolation metric"
network "percolation indicator"
network "percolation measure"
network "percolation index"
network "resilience metric"
network "resilience indicator"
network "resilience measure"
network "resilience index"
```

7.2. Tabla comparativa de categorías

Tabla 7.1: Tabla comparativa de categorías por artículo.

| Artículo | [7] | [87] | [63] | [65] | [62] | [61] | [69] | [87] | [66] | [67] | [59] | [91] | [57] | [33] | [58] | [42] | [43] | [78] | [14] | [93] |
|---------------------|-----|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| <i>Connectivity</i> | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | | | | | | | | ✓ | | ✓ | ✓ |
| <i>Degree</i> | | | | | | | | | | | | | | ✓ | | ✓ | | | | |
| <i>Percolation</i> | | | | | | | | | | | ✓ | | | | | | | | | |
| <i>Centrality</i> | | | | | | | | | ✓ | | ✓ | | | | ✓ | | | ✓ | | |
| <i>Distance</i> | | ✓ | | | | | | ✓ | | ✓ | | ✓ | ✓ | | | | | | | ✓ |
| <i>Spectral</i> | | | | | | | | | | | | | | | | | | | | |
| <i>Varies</i> | ✓ | | | | | | | | | | | | | | | | | | | |

Tabla 7.2: Tabla comparativa de categorías por artículo.

| | | | | | | | | | | | | | | | | | | |
|---------------------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| Artículo | [60] | [75] | [13] | [12] | [19] | [50] | [85] | [88] | [22] | [89] | [82] | [45] | [86] | [90] | [76] | [80] | [47] | [64] |
| <i>Connectivity</i> | | ✓ | ✓ | ✓ | | | | | | | | | | ✓ | ✓ | ✓ | ✓ | |
| <i>Degree</i> | | | | | ✓ | | | | | | ✓ | | | | | | | ✓ |
| <i>Percolation</i> | | | | | | | | | | ✓ | | | | | | | | |
| <i>Centrality</i> | ✓ | | | | ✓ | | | | ✓ | | | | ✓ | | | | | |
| <i>Distance</i> | | | | | | | | | | | | ✓ | | | | | | |
| <i>Spectral</i> | | | | | | | ✓ | ✓ | ✓ | | | | | | | | | |
| <i>Varies</i> | | | | | | | | | | | | | | | | | | |