Resilient Multiculture Network Design in the Presence of Exploit-Triggered Correlated Failures

N. Boettcher, Y. Prieto, S. E. Restrepo, and J. E. Pezoa, *Member*, IEEE

***Abstract*— Data networks are typically equipped with the same hardware and software stacks. Correlated attacks exploiting shared vulnerabilities at the nodes may result in massive failures that disrupt network operation. In this paper, multiple correlated failures that may negatively impact a monoculture network are analyzed and a methodology to reduce their effects is proposed. The proposed methodology consists of introducing diversity into the network components by optimally selecting both the number of different network nodes and their locations within the network. First, an algorithm is proposed to introduce node diversity in the topology considering nodes' vulnerability indexes, which are associated with node vendors. Next, two different optimal node placement algorithms are proposed. The first algorithm aims to cluster nodes of the same type to maintain network connectivity, while the second seeks to maximize the network centrality metric to identify key nodes in the network. Our results show that reliability can increase up to 50% when compared to a monoculture design.**

***Keywords*—Resilience, exploit, correlated failures, srng, diversity.**

I. INTRODUCCIÓN

La resiliencia a fallas en redes es un concepto fundamental en la actualidad para el diseño topológico y de mecanismos de ruteo, que ofrezcan servicios rápidos de recuperación y con la menor pérdida de datos posible. Además, la resiliencia permite garantizar que la mayor parte de la red se mantenga conectada después de un evento de falla. En los últimos años las fallas múltiples correlacionadas han llamado la atención de investigadores [1]-[5]. Aunque no son frecuentes, dada su magnitud, pueden causar enormes daños. Este tipo de falla puede surgir de fenómenos naturales o ser provocada por el hombre, y puede tener características geográficas asociadas a terremotos, huracanes, ataques con armas de destrucción masiva, entre otros. También puede ocurrir que varios elementos de la red compartan recursos comunes, por lo que cuando esos recursos fallan, se interrumpe el funcionamiento de los elementos de la red, ya sean dispositivos o enlaces [6], [7].

Otro tipo de fallas correlacionadas, menos estudiadas, provienen de las *monocultures*, es decir, cuando todos los componentes de la red pertenecen al mismo proveedor de hardware y poseen el mismo software. En [8] mencionan que "…un gran ataque coordinado contra Microsoft Windows o Cisco IOS tendría consecuencias desastrosas en Internet, dada la cuota de mercado dominante de estos sistemas finales y plataformas de enrutadores". En consecuencia, el hecho de tener en la red dispositivos que comparten riesgos del mismo tipo (SRNG por sus siglas en inglés) podría llevar a un riesgo enorme [9]: si se explota una vulnerabilidad existente en todos los dispositivos de la red, toda la red podría desmoronarse.

Generalmente, cuando los arquitectos diseñan una red, por razones prácticas, adquieren casi todos los enrutadores/conmutadores del mismo proveedor, basándose en el precio, el soporte técnico o el historial de la marca. Este proceder tiene ventajas tales como facilitar la configuración de los equipos de red y la comunicación entre los elementos de ésta. Sin embargo, con la diversidad de dispositivos de distintas clases, la correlación entre ellos se reduce cuando se enfrentan a ataques como *exploits* (vulnerabilidades 0-day) o asociados a especificaciones de hardware, aumentando así la resiliencia.

La *multiculture* o diversidad en los componentes de la red consiste en proporcionar alternativas, de modo que si se ataca un tipo particular de componente, no todos fallarán. Así podría decirse que cada proveedor representa una clase diferente, y la diversidad consiste en diferencias de kernel, componentes de hardware, protocolos, etc. Algunos ejemplos de vulnerabilidad son los componentes de hardware que pueden hacer que un componente de red falle [10] y una gran cantidad de vulnerabilidades de denegación de servicio reportadas por Common Vulnerabilities and Exposures [11]. La existencia de un solo tipo de dispositivo o poca diversidad implica un riesgo enorme debido a la alta correlación que existe entre equipos del mismo tipo ante fallas por vulnerabilidades de software comunes a estos. La red es comprometida masivamente por agresores que solo necesitarían un método de ataque para inhabilitar un equipo o un S.O. Mientras más diversa la red, más recursos debe invertir el atacante para identificar vulnerabilidades presentes en los elementos de la red y diseñar un plan para atacarlos a todos [12].

Este trabajo es una extensión de nuestro trabajo [13], y está enfocado en la búsqueda de estrategias de diseño de redes *multiculture* para aumentar la resiliencia de la red ante fallas múltiples correlacionadas, las que son generadas por *exploits*. Para el modelamiento de la red de datos se utilizará teoría de grafos. Los nodos estarán asociados a las tecnologías de enrutadores y la correlación entre las fallas estará dada por la pertenencia de los nodos a una de éstas, cuyo software presenta vulnerabilidades propensas a *exploits*. Las fallas se modelarán de forma tal que al ser afectado un proveedor, todos los nodos pertenecientes a éste fallarán y los demás

N. Boettcher, Departamento de Ingeniería Eléctrica, Universidad de Escuela de Informática y Telecomunicaciones, Universidad Diego Portales, Chile, nboettcher@udec.cl.

Y. Prieto, Departamento de Ingeniería Eléctrica, Universidad de Concepción, Concepción, Chile, yprietoh@udec.cl.

S. E. Restrepo, Departamento de Ingeniería Eléctrica, Universidad de Concepción, Concepción, Chile, and Departamento de Medio Ambiente y Energía, Universidad Católica de la Santísima Concepción, Concepción, Chile, srestrepo@udec.cl, srestrepo@ucsc.cl.

J. E. Pezoa, Departamento de Ingeniería Eléctrica, Universidad de Concepción, Concepción, Chile, jpezoa@udec.cl.

Corresponding author: Jorge E. Pezoa

permanecerán operativos. El diseño de la red se planteará como problemas de optimización entera, que especificarán dónde deben ser ubicados los nodos de cada proveedor en la red. Para el diseño se utilizaron dos métricas: agrupamiento y centralidad. La métrica de agrupamiento permite generar una mayor cantidad de componentes conectados en la red, que se traduce en mayor resiliencia ya que evita islas (grafos no conexos) post falla. La centralidad, por otra parte, implica mayor interconexión entre los nodos, lo cual deriva en la robustez de la red. Con esta metodología de diseño resiliente se evitará tener un *monoculture* en la red, garantizando que ésta asimile las fallas y, en cierta medida, continúe funcionando. Para analizar las estrategias de diseño propuestas se utilizará el ATTR como medida de comparación frente a redes donde la ubicación de nodos es totalmente aleatoria.

Este manuscrito se organiza de la siguiente manera. La Sección II presenta el estado del arte en *mono* y *multiculture* aplicados a resiliencia de redes. En la Sección III se define formalmente el problema abordado en este trabajo. En las Secciones IV y V se presentan los problemas de optimización de búsqueda de diversidad y de ubicación óptima de nodos diversos en la red. La Sección VI presenta los resultados del trabajo y las comparaciones realizadas mediante nueve topologías y la métrica de resiliencia ATTR. Finalmente, la Sección VII presenta las conclusiones del trabajo.

II. ESTADO DEL ARTE

La literatura ofrece varios trabajos de investigación y métodos para mejorar la resiliencia y la sobrevivencia siguiendo la idea de la diversidad en la red. En [12] se propone un nuevo paradigma para la sobrevivencia de la red basado en la heterogeneidad de ésta. Los investigadores afirmaron que los diferentes elementos de las redes que tienen la misma capacidad funcional son, en general, vulnerables a diferentes ataques de seguridad. Por lo tanto, una red que contiene elementos heterogéneos ofrece una mayor supervivencia a los ataques de seguridad en comparación a una red homogénea. Los autores introdujeron la idea de un espacio de diversidad, donde las dimensiones se asignan a diferentes capacidades funcionales de la red tales como: sistemas operativos, medios de comunicación, modelos de servicio, protocolos de red, mecanismos de enrutamiento, etc. Utilizando esta definición de espacio, los autores calcularon la distancia entre los elementos de la red y concluyeron que cuanto mayor sea la distancia entre pares de elementos de la red, mayor será la diversidad y menor la vulnerabilidad.

En la misma línea de razonamiento, en [14] se describió un algoritmo para aumentar la diversificación en la trayectoria de ataque de un adversario para alcanzar un activo. En [15] afirmaron que, si todas las rutas de ataque están perfectamente diversificadas, el adversario debe realizar un esfuerzo individual e independiente para explotar con éxito las vulnerabilidades que se encuentran en cada ruta de ataque. En [15], [16] se presentaron algoritmos para aumentar la conectividad en una red, basados en la diversidad de nodos y la forma en que se conectan entre sí. La idea principal de estos trabajos es que cuando ocurre un fallo, todos los nodos pertenecientes a la misma clase fallan simultáneamente. El método presentado en [15] divide la cantidad de nodos equitativamente entre las clases y crea una red en malla. Basado en la teoría de coloración de grafos se crean enlaces entre los nodos para que cada nodo de una clase esté conectado a nodos de clases diferentes. Este método de particionamiento de grafos genera topologías que cuando una clase de nodos desaparece, el resto de la red permanece conectada. En [16], el concepto de diversidad se propone como una solución para aumentar la resiliencia de la red frente a defectos de software, errores y vulnerabilidades en los routers. Los autores discutieron dos enfoques para la ubicación de nodos basados en la teoría de coloración de grafos: agrupación de grafos y partición de grafos. Los enrutadores están coloreados en función de sus implementaciones y la conectividad se maximiza cuando se elimina cualquier color del grafo. Para mejorar aún más la fiabilidad, se asignan funciones dentro de la red, donde los enrutadores de acceso y los enrutadores de red troncal se diferencian por los métodos de coloración. El balance de nodos entre las clases es óptimo si estos se distribuyen equitativamente, lo cual es una diferencia con nuestro trabajo. En [13], dada una red en particular, donde ya existen enlaces, se quieren especificar enrutadores de diferentes clases a los nodos de la red, para aumentar la resiliencia de la misma, evitando la aparición de *monoculture* a partir del diseño multiculture. Se plantean dos problemas de optimización secuenciales para el diseño multiclase de enrutadores de la red. El primero permite establecer la cantidad de elementos de cada clase en proporción inversa a su índice de vulnerabilidad, teniendo en cuenta una restricción debido al presupuesto. La solución a este problema permite la diversificación de tecnologías utilizadas, evitando la aparición de *monoculture*. El segundo problema selecciona una distribución de nodos por clase y los ubica en la topología escogida. La idea detrás del método es que ante la falla de una clase de nodos, puedan permanecer operativos la mayor cantidad posible de enlaces. Una forma de garantizar esto es agrupar los nodos de una misma clase, de modo que el fallo de una clase tenga mínimo impacto en la desconexión de enlaces conectando nodos de diferentes clases. Comentamos además que las principales diferencias con estos tres últimos trabajos son: (i) a nuestro entender, la idea de un índice de vulnerabilidad para cuantificar la fiabilidad de los proveedores y la ubicación de estos en la red; y (ii) a diferencia de [15], [16] en nuestro trabajo, consideramos clases no homogéneas.

Finalmente, en [17] se llevó a cabo un estudio, a través de una encuesta, para entender el problema al que se enfrentan los administradores de red, sus causas potenciales y cómo diagnosticarlas. La encuesta incluyó respuestas de administradores de redes pequeñas, medianas y grandes. Las causas más comunes reveladas por el estudio fueron fallas en el hardware, errores en el software del enrutador/conmutador y fallas externas.

III. PLANTEAMIENTO DEL PROBLEMA

La topología de una red de datos es modelada en este trabajo como un grafo no direccionado $G = (V, L)$, donde $V = \{1,2, \dots, n\}$ es el conjunto de nodos y $L = \{(i, j) : \text{nodos } i \text{ y } j \text{ están conectados}\}$ es el conjunto de enlaces bidireccionales entre los nodos. Supongamos que, durante

una actualización tecnológica de todos los nodos de la red, tenemos a nuestra disposición varios proveedores de equipamiento entre los cuales se puede seleccionar los nodos. Denotamos aquí por $K = 1,2,\dots,\kappa$ el conjunto de clases de dispositivos de red (proveedores) disponibles durante la actualización. También, asumimos que cada clase de dispositivos considerados aquí son propensos a ser atacados de forma individual. Más concretamente, cada clase puede ser atacada de forma correlacionada, aprovechando un *exploit* particular asociado a la clase, y la consecuencia de los ataques es que todos los nodos de la red pertenecientes a dicha clase fallarán simultáneamente.

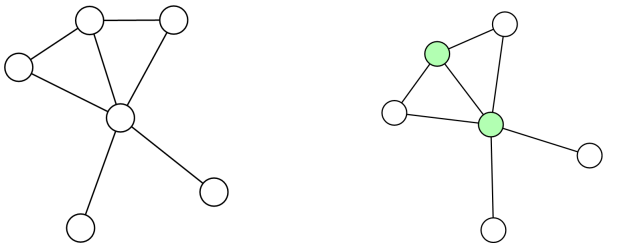


Figura 1. Comparación entre: (a) *monoculture* y (b) *multiculture* con ubicación aleatoria. Topología utilizada: Napnet.

Dado un grafo $G = (V, L)$ que representa la topología de una red con n nodos, y un total de κ diferentes clases de nodos de red, en este trabajo abordamos los siguientes problemas: (i) cuántos nodos de cada clase son necesarios para minimizar la vulnerabilidad de toda la red; y (ii) dónde deben ubicarse los nodos de cada clase para maximizar la conectividad media de la red. Para ilustrar que la solución no es trivial, consideremos la topología *monoculture* Napnet, de la Fig. 1a. En caso de un ataque a una vulnerabilidad de la clase, toda la red dejará de funcionar. Por su parte, la Fig. 1b muestra Napnet con nodos de dos clases distintas. Sin embargo, una incorrecta ubicación de los nodos compromete igualmente la red, ya que si una de las clases falla, la porción operativa podría quedar severamente desconectada.

IV. BÚSQUEDA DE DIVERSIDAD DE COMPONENTES EN LA RED

El primer problema de optimización abordado en este trabajo corresponde a especificar de forma óptima cuántos nodos de cada clase son necesarios para minimizar la vulnerabilidad de toda la red, cuando la topología no cambia y el arquitecto de red dispone de un presupuesto fijo para comprar nodos. Así el objetivo de este problema es forzar la diversidad para maximizar resiliencia. Este problema puede ser matemáticamente planteado como:

$$\mathbf{n}^* = \arg\min_{n \in S} \sum_{k=1}^{\kappa} \left(\alpha_k n_k - \alpha \frac{n}{\kappa}\right)^2, \quad (1)$$

sujeto a:

$$\sum_{k=1}^{\kappa} n_k B_k \le B \quad \text{(restricción de presupuesto)}, \quad (2)$$

$$\sum_{k=1}^{\kappa} n_k = n \quad \text{(restricción número total de nodos)}, \quad (3)$$

donde $\mathbf{n}^* = (n_1 \dots, n_k, \dots, n_\kappa)$ es un vector κ-dimensional que especifica el número óptimo de nodos de cada clase, $S \subset \mathbb{N}_0^\kappa$ es el espacio de búsqueda, B_k el costo económico de cada nodo de la k-ésima clase, B es el presupuesto total disponible para comprar nodos de red, y $\alpha_c = n^{-1} \sum_{k=1}^{\kappa} \alpha_k n_k$, como se demostró en [13]. El término α_k es un parámetro clave y lo denominamos *el índice de vulnerabilidad asociado al k-ésimo proveedor de equipamiento de nodos de red.*

Se hace notar que la función objetivo busca crear un equilibrio entre las clases dentro de la red, con una mayor presencia de nodos de clases menos vulnerables. Cuando las vulnerabilidades son idénticas, los términos $\left(n_k - \frac{n}{\kappa}\right)^2$ buscan que todas las clases tengan el mismo número de nodos. Por otra parte, los términos $\alpha_k n_k$ imponen que el número de nodos de una clase sea inversamente proporcional a su índice de vulnerabilidad, ya que el valor de α busca que el valor mínimo de la función objetivo se alcance en cero.

V. BÚSQUEDA DE UBICACIÓN DE COMPONENTES DIVERSOS EN LA RED

Después de determinar de forma óptima cuántos nodos de cada clase son necesarios para minimizar la vulnerabilidad de toda la red, se busca ahora asignar la ubicación de los nodos en la topología de red. Para esto se proponen dos métodos: uno que busca agrupar nodos de una misma clase y otro que ubica los nodos menos vulnerables en zonas claves de la red.

Agrupamiento de nodos de una misma clase

La idea de este algoritmo es ubicar los nodos en la topología de red de manera tal que, cuando todos los nodos de una misma clase fallen, la mayor cantidad posible de enlaces se mantengan operativos en la red. Para lograr esto, se propone acá agrupar los nodos de una misma clase en la topología de red. Por lo tanto, cuando una clase completa deje de funcionar, se busca que falle la cantidad mínima de enlaces que conectan los nodos de la clase fallida, con nodos de otras clases. Matemáticamente se define el problema de optimización entera con restricciones:

$$T^*(V) = \max_{T(V) \in \mathrm{T}} \sum_{k=1}^{\kappa} \frac{n_k}{N_k n} \sum_{(i,j) \in L^k} l_{ij}^k, \quad (4)$$

sujeto a:

$$\sum_{i=1}^{n} 1_{\{T(i)=k\}} = n_k, , \quad \text{(restricciones número de clases)}, \quad (5)$$

donde $k = 1, 2, \dots, \kappa$, $T(V): V \to K$ es un mapeo de V a K que asigna al i-ésimo nodo la clase $T(i) = k$, T es el espacio de búsqueda de todos los mapeos posibles para asignar las κ clases a los n nodos, $G^k = (V^k, L^k) \subset (V, L)$ es la topología de red resultante después de que una falla afecta a todos los nodos de la k-ésima clase, $l_{ij}^k = 1$ si $(i,j) \in L^k y l_{ij}^k = 0$ en caso

contrario, N_k es el número de componentes conectados en G^k, y $1_{\{T(i)=k\}} = 1$ es la función indicatriz de que el i-ésimo nodo pertenece a la clase k.

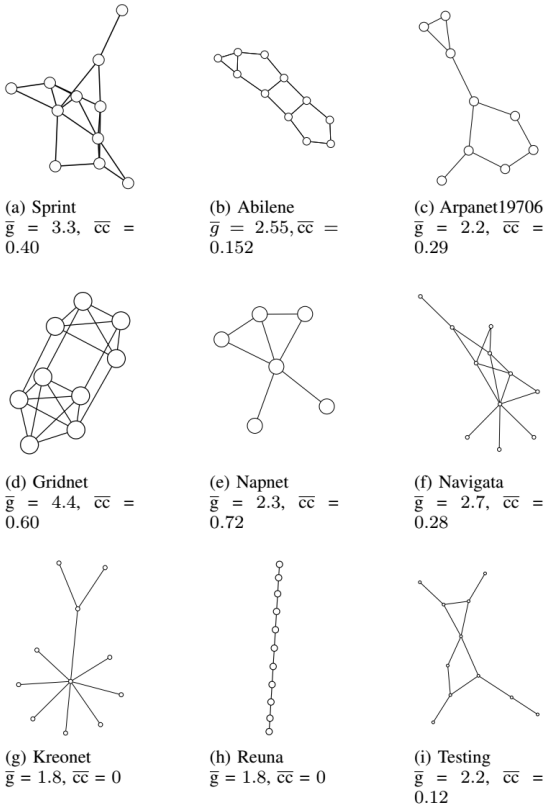


Figura 2. Topologías de redes utilizadas para la evaluación de los métodos de diversidad.

Notar que la función objetivo (4) busca penalizar todas aquellas soluciones (mapeos) que, después de la falla en todos los nodos de una misma clase, dejen nodos de otras clases desconectados en la red. Esto se observa claramente en el factor de penalización N_k, pues este valor será mínimo cuando la red resultante después del fallo de la k-ésima clase se mantenga totalmente conectada, es decir, el número de componentes conectados es $N_k = 1$, [18], [19]. En caso contrario, este valor aumenta en la medida que la red ya no es conexa y aparecen más y más componentes conectados, que están aislados entre sí, en la red [18], [19]. Por otra parte, las κ restricciones formuladas en (5) fuerzan la condición de los nodos de cada clase, que son necesarios para minimizar la vulnerabilidad de toda la red, que corresponde a la solución del problema de optimización (1)-(3). Finalmente, se hace notar que en [13] se observó que en redes que no están altamente conectadas, el ubicar nodos en la red de manera agrupada de acuerdo a la solución del problema de optimización (1)-(3), puede no ser suficiente para maximizar la resiliencia de la red. Esto se debe a que la falla de una clase completa de nodos puede implicar la desconexión de más elementos de otras clases. De ocurrir esto, es deseable que los nodos tomen posiciones dentro de la red de acuerdo a su índice de vulnerabilidad, para así asegurar la operatividad de la mayoría de estos en la red cuando las clases fallan.

Centralización de nodos en la red

La idea de este segundo algoritmo de ubicación de nodos en la red es seguir la siguiente intuición: "al diseñar una red, los nodos más robustos deben tomar posiciones de mayor importancia desde el punto de vista de conectividad". En ciencia de redes, una forma de medir la importancia de un nodo en una topología es a través de la centralidad por distancia [18], [19]. En este trabajo, mediante las definiciones entregadas en el Apéndice extendimos este concepto a clases de nodos. Así, se tiene entonces que aquellos nodos de las clases menos vulnerables debiesen ubicarse en posiciones más centralizadas de la red, mientras que los nodos más vulnerables debiesen ubicarse en la periferia. Con esto, el ubicar nodos de esta manera hace que la influencia de la falla de la clase menos (respectivamente, más) vulnerable sea mayor (respectivamente, menor) en mantener la conectividad completa de la red.

Así, se formula el siguiente problema de optimización entera con restricciones para ubicar los nodos en la red:

$$T^*(V) = \max_{T(V)\in \mathrm{T}} \overline{C} = \max_{T(V)\in \mathrm{T}} \sum_{k=1}^{\kappa} \alpha_k C(k), \qquad (6)$$

sujeto a:

$$\sum_{i=1}^{n} 1_{\{T(i)=k\}} = n_k, \quad \text{(restricciones número de clases)}, \quad (7)$$

donde $\overline{C}$ es la centralidad promediada de las clases de nodos de la red y $C(k)$ es la centralidad de la k-ésima clase.

VI. RESULTADOS

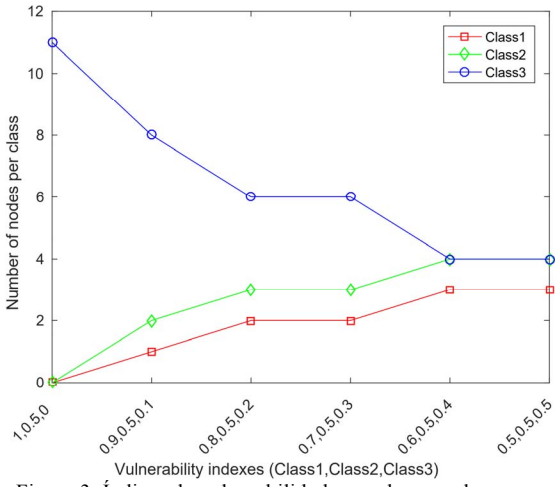


Figura 3. Índices de vulnerabilidad vs nodos por clases.

Topologías y métricas de resiliencia

En este trabajo se utilizan nueve topologías para evaluar los algoritmos propuestos. Estos se muestran en la Fig. 2 con sus respectivos grado medio ($\overline{g}$), abarcando redes de baja, mediana y alta conectividad; y sus coeficiente de agrupamiento medio ($\overline{cc}$). Las primeras ocho redes fueron extraídas de *Internet Topology Zoo* [20]. Las redes Navigata, Kreonet y Reuna (Fig. 2f, 2g, 2h respectivamente) son subgrafos de las redes originales. La red Testing (Fig. 2i) fue creada para probar nuestros algoritmos.

Para medir los resultados de los diseños propuestos se utiliza la fiabilidad promedio de dos terminales (ATTR por sus

siglas en inglés) [21], esta métrica ofrece información sobre qué tan bien conectada se mantiene la red después de los eventos de falla. Una versión adaptada de ATTR fue desarrollada en (9), donde ATTR(k) es el valor de ATTR después de que la clase k falla, y se define como:

$$\mathrm{ATTR}(k) = \binom{n}{2}^{-1} \sum_{i \neq j} Z_{ij}^{k}, \tag{8}$$

donde $\binom{n}{2}$ es el valor del coeficiente binomial y Z_{ij}^{k} es una variable binaria que toma el valor 1 en caso de existir un camino entre los nodos i y j después de una falla de la k-ésima clase, y toma el valor 0 en caso contrario. Luego, el ATTR de la red es calculado como la media ponderada de todas las clases que fallan, es decir:

$$\mathrm{ATTR} = \sum_{k=1}^{\kappa} \frac{\alpha_k}{\alpha} \mathrm{ATTR}(k), \tag{9}$$

donde $\alpha = \sum_{k \in K} \alpha_k$.

Diversidad en los nodos de la red

El primer resultado que se presenta en este trabajo es el efecto de los índices de vulnerabilidad en la cantidad de nodos por clase a especificar. La Fig. 3 muestra que, para tres clases de nodos y sin considerar la restricción presupuestaria (2), cuando los valores de los índices de vulnerabilidad en la tupla $(\alpha_1, \alpha_2, \alpha_3)$ son iguales o similares, la cantidad de nodos por clase resulta ser homogénea. De lo contrario, cuando se distancian los valores de los índices de vulnerabilidad, se seleccionan más nodos de la clase con índice menor.

En la Fig. 4 se puede observar la relación entre el presupuesto y la cantidad de nodos que se pueden adquirir para la red. A medida que se incrementa el presupuesto, es posible garantizar una mejor distribución de las clases de nodos, en función de sus índices de vulnerabilidad, ya que la restricción presupuestaria adquiere menos influencia en la función objetivo. (De ahora en adelante, los colores rojo, verde y azul representarán, respectivamente, clases con índices de vulnerabilidad 0.6, 0.5 y 0.4.)

El problema de búsqueda de diversidad en los componentes de la red, (1)-(3), es un problema NP-completo, que no considera la topología de la red, por lo que hay un compromiso entre diversidad y conectividad que se resuelve solucionando los problemas de ubicación óptima de nodos. Para mostrar la relevancia de la topología en una correcta elección de **n**, se utilizó un método de búsqueda exhaustivo de todas las combinaciones posibles de clases que satisfacen (2) y (3). Una vez obtenidos estos resultados, se seleccionaron, de entre el 10 % de los menores valores de la función de costo (1), todas aquellas soluciones factibles, **n**, que poseían el máximo valor de ATTR. Se utilizó sólo el 10 % de las soluciones factibles encontradas, ya que se observó que a medida que la selección de clases se aleja del óptimo de diversidad, éste tiende a priorizar la exclusividad de clases menos vulnerables, perdiendo de esta forma diversidad en la solución.

En la Tabla I, se listan los resultados de un método aleatorio de ubicación de nodos según las tuplas $\mathbf{n}^*$, el valor óptimo de selección de clases de acuerdo a (1)-(3), y **n** que pertenecen al 10 % mejor de soluciones factibles para caso de κ=3 clases. Para el método aleatorio se utilizó una función de densidad de probabilidad uniforme para asignar la ubicación de nodos en la red. La media y la desviación estándar de los resultados de ATTR se calcularon a partir de 100000 realizaciones aleatorias de ubicaciones de nodos. De la Tabla I se observa que el hecho de considerar una mayor cantidad de soluciones factibles que aquella que obtiene el óptimo en el problema (1)-(3) permite encontrar valores de ATTR mayores. Esto se explica puesto que el problema de optimización busca aumentar la diversidad de la red, considerando la vulnerabilidad de cada clase, pero no considera en su formulación la métrica ATTR. Con esto concluimos que, una buena heurística a utilizar para resolver el problema de optimización NP-completo de búsqueda de diversidad, consiste en obtener un conjunto de soluciones factibles en vez de obtener una única solución al problema.

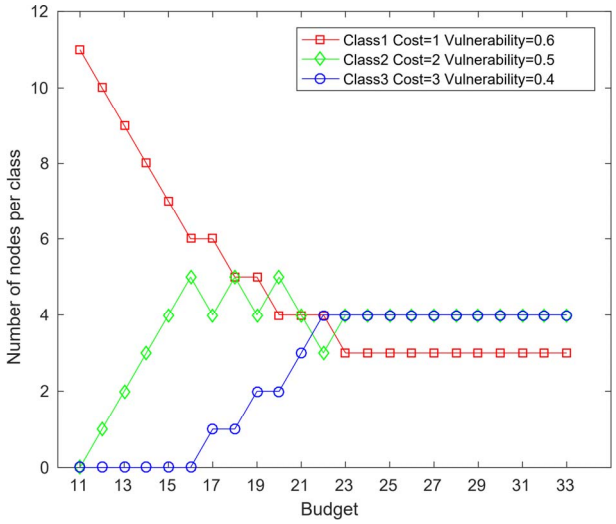
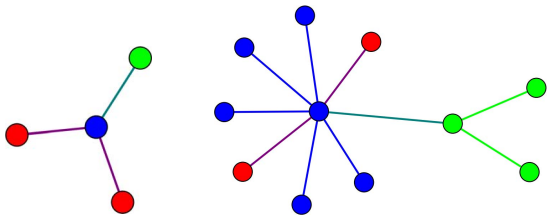


Figura 4. Presupuesto vs número de nodos por clase. Red con 11 nodos.

TABLA I
VALORES ATTR PARA $\mathbf{n}^*$ y **n** RESPECTIVAMENTE, PARA UBICACIÓN DE NODOS ALEATORIA.

Red	n	$\mathbf{n}^*$	$\overline{\mathrm{ATTR}} \pm \sigma$	**n**	$\overline{\mathrm{ATTR}} \pm \sigma$
Sprint	11	(3,4,4)	0.3638 ± 0.0411	(2,3,6)	0.4261 ± 0.0362
Abilene	11	(3,4,4)	0.3282 ± 0.0476	(2,3,6)	0.3726 ± 0.0500
Arpanet	9	(2,3,4)	0.3150 ± 0.0516	(2,2,5)	0.3286 ± 0.0545
Navigata	11	(3,4,4)	0.3192 ± 0.0515	(2,3,6)	0.3319 ± 0.0543
Kreonet	11	(3,4,4)	0.3047 ± 0.0501	(2,3,6)	0.3155 ± 0.0560
Testing	11	(3,4,4)	0.2900 ± 0.0491	(2,3,6)	0.2994 ± 0.0514
Gridnet	9	(2,3,4)	0.3121 ± 0.0441	(2,2,5)	0.3153 ± 0.0439
Napnet	6	(2,2,2)	0.3136 ± 0.0424	(1,2,3)	0.3220 ± 0.0454
Reuna	11	(3,4,4)	0.3020 ± 0.0448	(2,3,6)	0.3069 ± 0.0462

Por lo tanto, de aquí en adelante la solución al problema de optimización (1)-(3) consistirá en seleccionar un conjunto de valores factibles para el problema. Estas múltiples soluciones serán entonces utilizadas en los problemas de ubicación de nodos lo que permite reducir el espacio de búsqueda de soluciones.



(a) Kreonet lógico (b) Kreonet físico

Figura 5. Proceso de centralización de nodos de la misma clase. Cada nodo lógico representa un conjunto de nodos físicos de la misma clase.

Ubicación de nodos: agrupación vs centralización

A continuación se presentan y comparan los resultados obtenidos para los métodos de agrupación y centralización de nodos en la red, para las topologías de la Fig. 2, con $\kappa = 2$ $\kappa = 2$ y $\kappa = 3$ clases de nodos. Inicialmente se define un presupuesto B que resulta suficiente en todos los casos para alcanzar el mínimo de la función objetivo que garantiza la diversidad. Los índices de vulnerabilidad utilizados son (0.4,0.6) y (0.4,0.5,0.6) para $\kappa = 2$ y $\kappa = 3$, respectivamente.

Las Tablas II y IV listan los resultados obtenidos para el método de agrupamiento, con nodos de dos y tres clases respectivamente. Las Tablas III y IV listan los resultados para el método de centralización, con dos y tres clases respectivamente. A fin de realizar una mejor comparación, en las tablas también se listan los resultados de un método de ubicación de nodos aleatorio.

Como se puede observar de las Tablas II a V, los resultados de la métrica ATTR para tres clases son mejores que para sólo dos, esto porque el impacto de eliminar una clase es menos severo para la red cuando existe una red más diversa. De los resultados se observa además que cuando el grado medio de conectividad en la red es mayor, es más probable que los resultados de ubicación aleatoria se acerquen al valor óptimo de los algoritmos diseñados. Esto se explica porque cuanto más conectada esté la red, menos daño sufrirá por los *exploits* asociados a cada clase, independientemente de las posiciones que se les asigne a los nodos de cada clase, situación que resulta ser bastante intuitiva. Por otra parte, se observa también que la ubicación de nodos al azar en la topología de red no genera mayores valores para la métrica de resiliencia ATTR, lo que refuerza la afirmación hecha anteriormente de que la ubicación de nodos debe realizarse mediante alguna metodología de asignación.

Si se compara ahora los métodos de ubicación de nodos propuestos, de las Tablas se puede concluir que el método de ubicación por centralización obtiene mayores valores para la métrica de ATTR. Este mejor desempeño se atribuye al hecho de que el método de centralización considera las distancias entre los nodos, lo que se relaciona de manera directa con el ATTR. Por otra parte, el método de ubicación por centralidad considera la vulnerabilidad de cada clase, mientras que el método de agrupamiento no los considera, pues busca reducir el número de componentes desconectados en una red frente a fallas correlacionadas.

Hacemos notar además que el método de ubicación por centralización genera topologías que para efectos de análisis pueden agregarse fusionando nodos de una misma clase en un sólo nodo, tal como se observa en la Fig. 5. Más aún, hacemos notar que en el 78 % de las redes analizadas, la topología generada por agrupación se logra reducir a un cliqué de tamaño tres, obteniendo de esta manera el valor máximo de la función de costo (6) en la mayoría de las tuplas pertenecientes a **n** para $\kappa = 3$. Con esto el método de ubicación por centralización es capaz de relajar aún más la restricción de diversidad y otorgar tuplas con mayor ATTR. Por ejemplo, para la topología Reuna se logró reducir a un grafo de tres nodos, pero es imposible lograr un cliqué debido a su topología. El único caso donde el grafo agregado no logró reducirse a tres nodos fue en Kreonet para $\kappa = 3$, donde se obtiene el mismo **n** y ATTR para ambos métodos. Para $\kappa = 2$ también se obtuvieron cuatro topologías con el mismo **n** y ATTR para ambos métodos. En este caso, a diferencia del anterior, el conjunto de valores factibles **n** se reduce a sólo un resultado, dada la pequeña cantidad de combinaciones para generar diversidad.

La Fig. 6 ilustra las tres topologías con menor conectividad, donde se observan las ubicaciones resultantes para los nodos pertenecientes a éstas para $\mathbf{n} = (3,4,4)$. En las topologías Kreonet, Figs. 6a y 6b, y Reuna, Figs. 6e y 6f, es fácil observar cómo la posición de la clase con el índice de vulnerabilidad mínimo aumenta la resiliencia de la red. Para la topología Testing, Figs. 6c y 6d, es evidente que el método de agrupamiento asigna las ubicaciones considerando N_k, donde no centralizar las clases, en este caso en particular, otorga un mayor ATTR.

TABLA II
VALORES ATTR PARA UBICACIÓN POR AGRUPAMIENTO Y DE FORMA ALEATORIA CON $\kappa = 2$

Red	n	Tipo de ubicación de nodos	
		Agrupamento ATTR	Aleatória $\overline{\text{ATTR}} \pm \sigma$
Sprint	(5,6)	0.2364	0.1572± 0.0478
Abilene	(5,6)	0.2364	0.1130± 0.0515
Arpanet	(4,5)	0.2333	0.1024± 0.0528
Navigata	(5,6)	0.2364	0.1384± 0.0465
Kreonet	(4,7)	0.2509	0.1399± 0.0774
Testing	(5,6)	0.2364	0.0909± 0.0423
Gridnet	(4,5)	0.2333	0.2101± 0.0409
Napnet	(2,4)	0.2667	0.1934± 0.0786
Reuna	(5,6)	0.2364	0.0620± 0.0328

Aplicación: actualización tecnológica de red

Un escenario interesante de optimización para implementar la diversidad y aumentar la resiliencia se da durante la

actualización tecnológica en una red. En ese caso es deseable mantener los enlaces y simplemente reemplazar todos los nodos, buscando el aumento de: capacidad de procesamiento, manejabilidad, seguridad, resiliencia, etc. Se presenta entonces acá una metodología para aumentar la resiliencia en un escenario de actualización tecnológica de redes. En la Fig. 7 se presenta un diagrama de flujo que combina los problemas de diversidad de componentes y de ubicación de nodos.

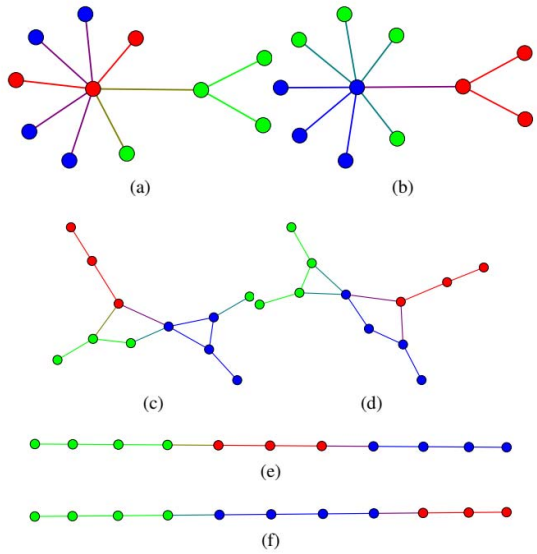


Figura 6. Topologías de menor conectividad (Kreonet, Testing y Reuna) donde (a), (c), (e) corresponden a ubicación por agrupamiento y (b), (d), (f) a ubicación por centralización.

Para saber cuántos nodos adquirir de cada clase, es necesario conocer el presupuesto de adquisición de nodos, el número de clases, el costo e índice de vulnerabilidad de cada dispositivo correspondiente a una clase, y el número total de nodos. La etapa de ubicación depende de la topología de red, porque los enlaces que ya existen serán reutilizados, además del uso o no de la información del índice de vulnerabilidad en el diseño.

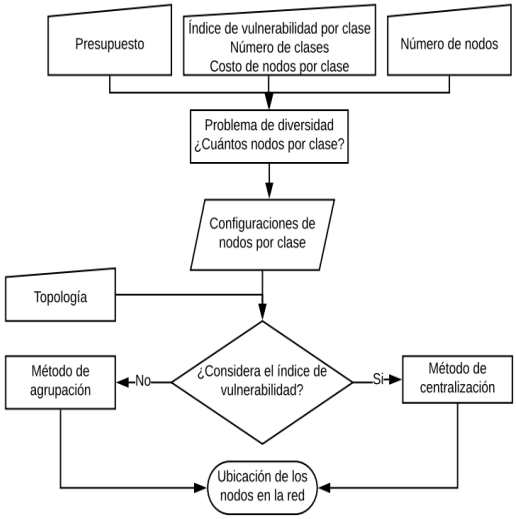


Figura 7. Proceso de actualización tecnológica de red.

IV. CONCLUSIONES

La diversidad como medio para proporcionar resiliencia resulta ser un factor clave en el diseño de redes. El método de optimización propuesto para calcular la cantidad de nodos por clase proporciona una metodología simple que permite distribuir la presencia de diferentes proveedores en la red, de acuerdo a los índices de vulnerabilidad, costos y presupuesto. Una vez que el arquitecto de red ha diversificado las clases a utilizar, las dos metodologías de ubicación de nodos en la red permiten aumentar la resiliencia de la red localizando de manera agrupada o centralizada los nodos, frente a fallas correlacionadas de las clases ocasionadas por *exploits*. Basados en la métrica ATTR, el método de ubicación por centralidad obtuvo mejores resultados pues permite equilibrar mejor la diversidad de clases y su vulnerabilidad, pues ambos factores son considerados durante el proceso de optimización.

TABLA III
VALORES ATTR PARA UBICACIÓN POR CENTRALIZACIÓN Y DE FORMA ALEATORIA CON $\kappa = 2$

Red	n	Tipo de ubicación de nodos: Agrupamento ATTR	Aleatória $\overline{ATTR} \pm \sigma$
Sprint	(4,7)	0.2727	0.2055± 0.0524
Abilene	(4,7)	0.2727	0.1555± 0.0616
Arpanet	(4,5)	0.2333	0.1037± 0.0527
Navigata	(4,7)	0.2727	0.1709± 0.0626
Kreonet	(4,7)	0.2509	0.1401± 0.0776
Testing	(4,7)	0.2727	0.1191± 0.0560
Gridnet	(4,5)	0.2333	0.2097± 0.0414
Napnet	(2,4)	0.2667	0.1936± 0.0786
Reuna	(4,7)	0.2727	0.0800± 0.0409

TABLA IV
VALORES ATTR PARA UBICACIÓN POR AGRUPAMIENTO Y DE FORMA ALEATORIA CON $\kappa = 3$

Red	n	Tipo de ubicación de nodos: Agrupamento ATTR	Aleatória $\overline{ATTR} \pm \sigma$
Sprint	(3,4,4)	0.4327	0.3632± 0.0411
Abilene	(3,4,4)	0.4327	0.2940± 0.0559
Arpanet	(2,3,4)	0.4463	0.2820± 0.0602
Navigata	(3,4,4)	0.4327	0.2916± 0.0417
Kreonet	(2,3,6)	0.4461	0.3012± 0.0729
Testing	(3,3,5)	0.4461	0.2286± 0.0457
Gridnet	(2,3,4)	0.4463	0.4415± 0.0131
Napnet	(1,2,3)	0.4533	0.3685± 0.0560
Reuna	(4,4,3)	0.3382	0.1331± 0.0362

TABLA V
VALORES ATTR PARA UBICACIÓN POR CENTRALIZACIÓN Y DE FORMA ALEATORIA CON $\kappa = 3$

Red	n	Tipo de ubicación de nodos: Agrupamento ATTR	Aleatória $\overline{ATTR} \pm \sigma$
Sprint	(2,3,6)	0.4800	0.4259± 0.0360

Abilene	(2,3,6)	0.4800	0.3815± 0.0604
Arpanet	(1,3,5)	0.4944	0.3663± 0.0669
Navigata	(2,3,6)	0.4800	0.3417± 0.0531
Kreonet	(2,3,6)	0.4461	0.2981± 0.0735
Testing	(2,3,6)	0.4800	0.2708± 0.0557
Gridnet	(1,3,5)	0.4944	0.4859± 0.0122
Napnet	(1,1,4)	0.5067	0.4412± 0.0743
Reuna	(2,3,6)	0.4509	0.1854± 0.0516

AGRADECIMIENTOS

Este trabajo fue apoyado por CONICYT: FONDECYT Regular 2016 Folio 1160559, PCHA/Doctorado Nacional Folios 2015-21150775 y 2015-21150313.

APÉNDICE A

En ciencia de redes, la centralidad por distancia se define como el inverso de la distancia media entre todos los nodos de red [18], [19]. De manera análoga, acá definimos, para la red $G = (V, L)$ que presenta el mapeo de n nodos a κ clases $T(V)$, la distancia media del nodo i, que pertenece a la clase l, a la k-ésima clase como: $D_i(k) = \frac{1}{n_k} \sum_{j \in \{T(j)=k\}} d_{i,j}$, donde $d_{i,j}$ es la distancia, en número de enlaces, entre los nodos i y j. Por lo tanto, la distancia media de la clase l a la clase k se define como: $D(l,k) = \frac{1}{n_l} \sum_{j \in \{T(j)=l\}}^{n_l} D_j(k)$. Así, la centralidad de la k-ésima clase se define como:

$$C(k) = \frac{1}{\sum_{j=1}^{n_l} D(k,j)} \tag{12}$$

REFERENCIAS

[1] Diaz, O., Xu, F., Min-Allah, N., Khodeir, M., Peng, M., Khan, S., and Ghani, N. Network survivability for multiple probabilistic failures. IEEE Communications Letters, 16(8), 1320-1323. 2012.

[2] Lee, H. W., Modiano, E., and Lee, K. Diverse routing in networks with probabilistic failures. IEEE/ACM TON, 18(6), 1895-1907. 2010.

[3] Izaddoost, A., and Heydari, S. S. Enhancing network service survivability in large-scale failure scenarios. Journal of Communications and Networks, 16(5), 534-547. 2014.

[4] Rahnamay-Naeini, M., Pezoa, J. E., Azar, G., Ghani, N., and Hayat, M. M. Modeling stochastic correlated failures and their effects on network reliability. In Proc. ICCNC 2011 (pp. 1-6). IEEE. 2011.

[5] Neumayer, S., and Modiano, E. Network reliability under random circular cuts. In Proc. GLOBECOM 2011, (pp. 1-6). IEEE. 2011.

[6] Freitas R. et al., "OSNR-based Restoration Algorithm for Optical Network Resilience to Node Failures," in IEEE Latin America Transactions, vol. 10, no. 4, pp. 1893-1900, 2012.

[7] Paez V., Talia A., Davalos E. and Pinto D., "Optimal Selection of p-cycles on WDM Optical Networks with Shared Risk Link Group Independent Restorability using Genetic Algorithm," in IEEE Latin America Transactions, vol. 10, no. 1, pp. 1385-1390, 2012.

[8] Sterbenz, J. P., Hutchison, D., Çetinkaya, E. K., Jabbar, A., Rohrer, J. P., Schöller, M., and Smith, P. Redundancy, diversity, and connectivity to achieve multilevel network resilience, survivability, and disruption tolerance. Telecommunication Systems, 56(1), 17-31. 2014.

[9] Iqbal, F., and Kuipers, F. A. Disjoint paths in networks. Wiley Encyclopedia of Electrical and Electronics Engineering. 2015.

[10] Cisco. Clock Signal Component Issue. Access Date: 7th June 2017. Available in: http://www.cisco.com/c/en/us/support/web/clock-signal.html#~overview

[11] CVE. Vulnerability Trends Over Time. Access Date: 7th June 2017. Available in: http://www.cvedetails.com/product/19/Cisco-IOS.html?vendor_id=16

[12] Zhang, Y., Vin, H., Alvisi, L., Lee, W., and Dao, S. K.. Heterogeneous networking: a new survivability paradigm. In Proc. Workshop on New Security Paradigms (pp. 33-39). ACM. 2001.

[13] Prieto, Y., Pezoa, J. E., Boettcher, N., and Sobarzo, S. K. Increasing network reliability to correlated failures through optimal multiculture design. In Proc. Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON), 2017 CHILEAN Conference on (pp. 1-6). ISSN: 0719-6806, IEEE. 2017.

[14] Bopche, G. S., and Mehtre, B. M. Exploiting curse of diversity for improved network security. In Proc. ICACCI 2015 (pp. 1975-1981). IEEE. 2015.

[15] Zhu, Y., and Huang, X. Node robust algorithm study based on graph theory. In Proc. Fuzzy Systems and Knowledge Discovery (FSKD) 2011 (vol. 4, pp. 2300-2303). IEEE. 2011.

[16] Caballero, J., Kampouris, T., Song, D., and Wang, J. Would diversity really increase the robustness of the routing infrastructure against software defects?. Department of Electrical and Computing Engineering, 40, 2008.

[17] Zeng, H., Kazemian, P., Varghese, G., and McKeown, N. A survey on network troubleshooting. Technical Report Stanford/TR12-HPNG-061012, Stanford University, Tech. Rep.

[18] Newman, M., Barabási, A.-L., and Watts, D. J. The Structure and Dynamics of Networks. The Princeton Press, 2006.

[19] Lewis, T. G. Network Science: Theory and Applications. Wiley, March 11, 2009.

[20] Zoo, T. Access Date: 7th June 2017. Available in: http://www.topology-zoo.org

[21] Rai, S., and Agrawal, D. P. Distributed computing network reliability. 1990.

Nicolás Boettcher Recibió los grados de B.Sc. en Ingeniería Civil en Informática y Telecomunicaciones y M. S. de la Ingeniería con mención en Informática y Telecomunicaciones en 2008 y 2015, respectivamente, de la Universidad Diego Portales, Santiago, Chile. Además, es Ph.D. (c) en Ciencias de la Ingeniería con mención en Ingeniería Eléctrica de la Universidad de Concepción. Actualmente es académico en la Escuela de Informática y Telecomunicaciones de la Universidad Diego Portales, Santiago, Chile. Sus áreas de interés son análisis de protocolos de red, medición y simulación, seguridad y privacidad de la información y reconocimiento de patrones.

Yasmany Prieto Recibió el grado de B. Sc. en Ingeniería en Telecomunicaciones y Electrónica en 2012 en la Universidad de Pinar del Río, Pinar del Rio, Cuba. Actualmente es estudiante de Doctorado en el programa de Ciencias de la Ingeniería con mención en Ingeniería Eléctrica de la Universidad de Concepción, Concepción, Chile. Sus áreas de interés son optimización de redes, reconocimiento de patrones, procesamiento de señales.

Silvia Elena Restrepo Recibió los grados de B.Sc. en Ingeniería Física y M. S. en Ingeniería de Sistemas en 2009 y 2012, respectivamente, de la Universidad Nacional de Colombia, Sede Medellín, Colombia. Además, recibió el grado de Ph. D en Ciencias de la Ingeniería con mención en Ingeniería Eléctrica de la Universidad de Concepción, Chile, en 2016. La Dra. Restrepo actualmente es Profesora Auxiliar del Departamento de Medio Ambiente y Energía, Universidad Católica de la Santísima Concepción, Concepción, Chile. Sus áreas de interés son Redes de Sensores Inalámbricas, Sistemas Distribuidos, Sistemas Inteligentes e Inteligencia Ambiental.

Jorge E. Pezoa (S'08-M'10) recibió los grados de B.Sc. y M. S. en Ingeniería Electrónica e Ingeniería Eléctrica en 1999 y 2003, respectivamente, de la Universidad de Concepción, Concepción, Chile. Además, recibió el grado de Ph.D. en Ingeniería de University of New Mexico en

2010. El Dr. Pezoa actualmente es Profesor Asociado y Subdirector del Departamento de Ingeniería Eléctrica de la Universidad de Concepción, en Concepción, Chile, y es miembro de Society of Photo-optical Instrumentation Engineers (SPIE), Optical Society of America (OSA), y Association for Computing Machinery (ACM). Sus áreas de interés son procesamiento distribuido de información, reconocimiento de patrones, procesamiento estadístico de señales, optimización de redes, procesamiento de señales e imágenes hiperespectrales e infrarrojas para procesos industriales.