

Criterio previsor para la selección de pares en redes P2P móviles

Trabajo final de graduación
para optar por el grado de

Magister scientiæ en Telemática

Autor
Andrés Meseguer Rojas

Asesor
Francisco J. Torres Rojas, Ph.D.

noviembre 2015



Dedico este trabajo a mi madre, Floribeth,
quien me enseñó a buscar la felicidad.

Agradecimientos

Primeramente quiero agradecerle a usted, lector, por tomar su tiempo para consultar este texto. Espero que esta obra contenga el conocimiento que usted busca, o al menos pueda servirle de referente al lugar donde están las respuestas que necesita.

A Mariela Romero Zúñiga, por su finísimo trabajo, pues no solo hizo la revisión filológica, sino que fue fuente de apoyo para sacar adelante la investigación.

Finalmente, quiero agradecer a Francisco J. Torres, Ricardo Román y a todo el grupo de tesarios, los *Happy Few*, por ayudarme a ver este trabajo más allá de un requisito para una institución, sino como un ejercicio en búsqueda del conocimiento.

Resumen

La estructura de telefonía tradicional con torres que centralizan el flujo tiene una oferta inferior al nivel de demanda creciente. Enfoques P2P han sido exitosos en generar avances hacia una solución.

Sin embargo, la movilidad de los dispositivos supone un reto donde su patrón de movimiento es difícil de prever y está determinado por la voluntad de seres vivos.

De esta manera, se presenta un algoritmo predictor que, mediante el uso del tiempo de conexión y la señal de sus vecinos como entrada permita seleccionar el mejor par para ingresar a una red P2P.

Con el uso de un simulador multiagente se desplegó el algoritmo sobre escenarios de convergencia/divergencia, concentración y escuadrón para comparar sus resultados sobre la cantidad de desconexiones contra los criterios de redes convencionales.

Mediante un análisis de varianza se determinó que el criterio predictor es superior en redes medianamente densas. Entonces, la evidencia estadística sugiere que el algoritmo puede minimizar la cantidad de desconexiones en una red P2P.

Palabras clave: P2P, telefonía móvil, criterio predictor, movilidad.

Abstract

The traditional telephony structure, with towers centralizing the communication is offering a service unable to match the on-growing demand. P2P approaches have made progress towards a solution.

However, the devices's mobility represents a challenge where their migration pattern is difficult to predict and is determined by the will of living beings.

A foresighted algorithm, using the connection time and signal from their neighbors as input and selecting the best peer allowing access to a P2P network is presented.

Using a multi-agent simulation, the algorithm was executed under convergence/divergence, concentration and squad scenarios in order to compare their results on the number of disconnections against the criteria of conventional networks.

Through an analysis of variance, it was determined that the foresighted approach is superior in fairly dense networks. Statistical evidence suggests then that the algorithm can minimize the amount of disconnections in a P2P network.

Key words: P2P, mobile, foresighted, mobility.

Contenido

1	Introducción	1
1.1	Hipótesis	3
1.2	Objetivos	3
1.2.1	Objetivo general	3
1.2.2	Objetivos específicos	4
1.3	Alcances y limitaciones	4
2	Marco teórico y trabajo relacionado	7
2.1	¿Qué es una red P2P?	10
2.1.1	Teorema de CAP	11
2.1.2	Despliegue de una red P2P	14
2.1.3	Enrutamiento en una red P2P	15
2.2	Tecnologías móviles e inalámbricas	15
2.3	Algoritmos con criterio predictor	16
2.3.1	Criterios de selección	18
2.4	Otras propuestas en el campo	21

2.5	Importancia	24
3	Criterio previsor: desarrollo y análisis	26
3.1	Simulador	27
3.1.1	RVO2	29
3.1.2	Escenarios	30
3.1.3	Patrones de movimiento	30
3.2	Desarrollo del criterio previsor	37
3.2.1	Función de puntuación	39
3.3	Análisis de complejidad	42
3.3.1	Análisis de complejidad espacial	43
3.3.2	Análisis de complejidad temporal	43
4	Diseño de los experimentos	47
4.1	Factores y niveles	49
4.2	Variables de respuesta	51
4.3	Recolección de datos	51
4.4	Análisis de varianza	51
5	Resultados	54
5.1	Experimentos	55
5.2	Supuestos de ANOVA	55
5.2.1	Transformaciones Box-Cox	58
5.2.2	ANOVA sobre una distribución Weibull	61
5.3	Análisis de varianza	62
5.3.1	Factor algoritmo	65
5.3.2	Factor tamaño	66
5.3.3	Factor pasos	67
5.3.4	Factores algoritmo y tamaño	68

5.3.5	Factores migración y tamaño	69
5.3.6	Factores algoritmo y pasos	71
5.3.7	Factores tamaño y pasos	72
5.3.8	Factores algoritmo, tamaño y pasos	73
5.3.9	Factores migración, tamaño y pasos	74
5.3.10	Resumen	75
6	Conclusiones y trabajo futuro	76
6.1	Discusión	77
6.1.1	El criterio previsor	77
6.1.2	Selección de los factores	77
6.1.3	Redes densas	78
6.2	Trabajo futuro	79
6.3	Conclusiones	80
A	Creación de corridas aleatorias	81
B	Creación de escenarios aleatorios	83
C	Ejemplo de un escenario CD	89
D	Acerca de PEDSIM	91
E	Muestra de los resultados del simulador	93
	Bibliografía	94
	Licencia	100
F.1	Creative Commons Atribución – Compartir Igual 3.0 Costa Rica	100

Lista de figuras

1.1	¿Es posible una red móvil descentralizada?	3
2.1	Eventos como el terremoto de Cinchona, en Costa Rica, pueden desestabilizar las telecomunicaciones. Fuente: Wikimedia [5].	9
2.2	Abonados a telefonía fija o móvil por cada 100 personas. Fuente: Oficina del Informe sobre Desarrollo Humano [16].	9
2.3	La intensidad de la señal de un dispositivo móvil depende de la ubicación de este en un punto determinado del tiempo.	20
2.4	Seleccionar a un nodo solo por tener una señal más fuerte no necesariamente es el mejor criterio, pues esta podría ser poco duradera.	21
2.5	En un modelo híbrido, el fallo de una estación implica la desconexión total de los dispositivos. Estos no solo quedan estos aislados de la red, sino también imposibilitados de comunicación alguna entre ellos.	22
2.6	En un modelo híbrido, la movilidad de un dispositivo se ve reducida por el alcance de su estación.	23

2.7	Los escenarios más beneficiados por la investigación pueden ser: a) redes convencionales congestionadas durante desastres naturales o eventos sociales b) equipos de rescate c) poblaciones humanas alejadas de las redes móviles convencionales.	25
3.1	Ejemplo de la ejecución de un escenario con patrón de convergencia/divergencia en tres pasos distintos.	32
3.2	Ejemplo de la ejecución de un escenario con patrón de concentración en pasos distintos.	34
3.3	Ejemplo de la ejecución de un escenario con patrón de escuadrón en pasos distintos.	35
3.4	Ejemplo de la ejecución de un escenario con patrón aleatorio en pasos distintos.	36
3.5	Gompertz es una función sigmoide.	40
3.6	Proyección de las funciones de puntaje para el tiempo (tm_score) y para la señal (sg_score) usadas en la función global de puntaje previsor. X representa la intensidad del parámetro de entrada, mientras que Y representa el puntaje.	41
3.7	Proyección de la función de puntaje previsor. Su resultado varía de acuerdo con los parámetros de tiempo (time) y señal (signal).	42
5.1	Histograma para la frecuencia de las desconexiones.	57
5.2	Gráfico Q-Q para el análisis de normalidad en la varianza de las desconexiones.	57
5.3	Gráfico de ubicación escala para la varianza de las desconexiones.	57
5.4	Histograma para la frecuencia del logaritmo de las desconexiones.	60
5.5	Gráfico Q-Q para el análisis de normalidad en la varianza del logaritmo de las desconexiones.	60
5.6	Gráfico de ubicación escala para la varianza del logaritmo las desconexiones.	60

5.8	Resultados de la función <i>fitdist</i> para verificar que los datos pueden ser procesados por ANOVA. Con que los resultados sean mayores o cercanos a 0.5 se puede concluir que siguen una distribución Weibull.	61
5.7	Herramientas gráficas para la valoración visual de la normalidad en la variable de respuesta bajo una distribución Weibull. En la imagen se aplican a la varianza del logaritmo de las desconexiones.	62
5.9	Representación gráfica de los efectos del factor algoritmo sobre el logaritmo natural de las desconexiones.	64
5.10	Representación gráfica de los efectos del factor tamaño sobre el logaritmo natural de las desconexiones.	66
5.11	Representación gráfica de los efectos del factor tamaño sobre el logaritmo natural de las desconexiones.	67
5.12	Análisis de efectos para los factores algoritmo y tamaño. Significancia sobre el 99.9%.	68
5.13	Análisis de efectos para los factores migración y tamaño. Significancia sobre el 99.9 %.	69
5.14	Comportamientos interesantes al tratar poblaciones de agentes muy densas.	70
5.15	Análisis de efectos para los factores algoritmo y pasos. Significancia sobre el 99.9 %.	71
5.16	Análisis de efectos para los factores tamaño de la red, patrón de migración, algoritmo y tiempo de la vida de la red. Significancia sobre el 99.9 %.	72
5.17	Interacción de los factores algoritmo y tiempo para poblaciones de tamaño nivel 1.	73
5.18	Interacción de los factores algoritmo y tiempo para poblaciones de tamaño nivel 2.	73
5.19	Interacción de los factores algoritmo y tiempo para poblaciones de tamaño nivel 3.	73

5.20	Interacción de los factores migración y tiempo para poblaciones de tamaño nivel 1.	74
5.21	Interacción de los factores migración y tiempo para poblaciones de tamaño nivel 2.	74
5.22	Interacción de los factores migración y tiempo para poblaciones de tamaño nivel 3.	74

Introducción

“In 2014, the mobile industry contributed a total of US\$3 trillion to the world economy in value added terms, equivalent to around 3.8 % of global GDP”.

GSMA Global Mobile Economy
Report 2015, GSMA Intelligence.

Este trabajo constituye una investigación por parte del estudiante Andrés Meseguer Rojas para aplicar por el grado de maestría en Ingeniería en Computación con enfoque de Telemática, en el Instituto Tecnológico de Costa Rica.

Se busca la creación de redes de pares (P2P)¹ móviles cuyo tiempo de despliegue sea mínimo y las relaciones entre sus componentes sean duraderas y formadas con base en la búsqueda del máximo beneficio (figura 1.1).

El principal reto para generar estas relaciones es el constante cambio y movimiento de los agentes dentro de la red. Modelar su comportamiento no es sencillo, ni mucho menos anticiparlo.

Esta situación genera muchas desconexiones que necesitan ser tratadas con algún tipo de mecanismo de recuperación; no obstante, el tiempo que requieren y su efecto principal en el servicio que ofrece la red hace que se busque disparar este evento lo menos posible.

De esta manera, un algoritmo previsor establece un puntaje con base en el tiempo e intensidad de la señal de sus pares, lo cual permite minimizar las desconexiones, y mejorar así su desempeño.

Así, mediante el uso de un simulador multiagente se generaron 3840 diferentes escenarios, cada uno distinto en cuanto a su densidad de agentes, tiempo de ejecución y el patrón de migración aplicado.

Tras recopilar los resultados de estas simulaciones, se procedió a su estudio de acuerdo con una metodología de análisis de varianza. Este proceso reveló que el criterio previsor muestra sus mejores resultados en redes medianamente densas, y que justamente la densidad muestra un efecto de consideración especial.

¹Si bien “redes de pares” es una traducción correcta, las siglas P2P del inglés *peer-to-peer* tienen mayor aceptación y por ello serán usadas en este documento.

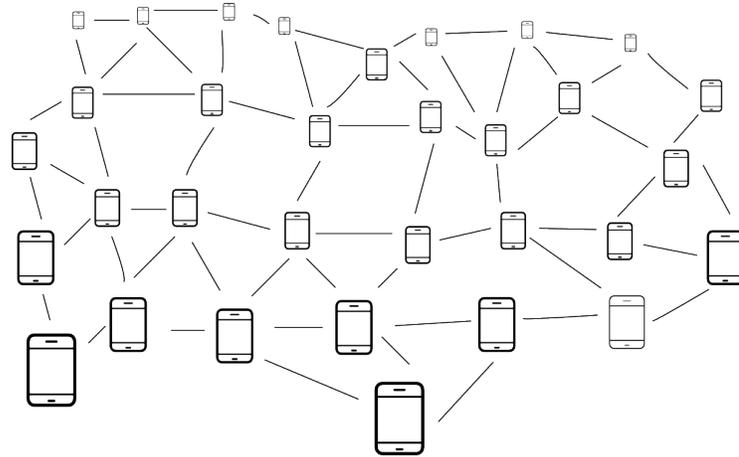


Figura 1.1: ¿Es posible una red móvil descentralizada?

1.1 Hipótesis

Al necesitar una red P2P cuyas relaciones sean duraderas y de máximo beneficio, se propone que:

Con el uso de un criterio predictor para la selección de pares en redes P2P móviles se obtiene una menor cantidad de desconexiones que al utilizar los criterios de señal, tiempo o azar por separado.

Así, es importante poder desplegar esta red en situaciones de rescate, congestión y de inexistencia de la estructura tradicional de telefonía móvil por fallo o inaccesibilidad.

1.2 Objetivos

1.2.1 Objetivo general

Minimizar el retraso y pérdida de datos que generan las constantes entradas, salidas y movimientos de los nodos en una red P2P, mediante el uso de un algoritmo con criterio

previsor.

1.2.2 Objetivos específicos

1. Minimizar la cantidad de ejecuciones de un mecanismo de recuperación de errores y el tiempo que este requiere.
2. Mejorar el desempeño al reducir la cantidad de reconexiones como resultado del uso de un criterio previsor para enrutamiento.
3. Desarrollar una red P2P cuyo tiempo de despliegue sea bajo.
4. Determinar cómo la movilidad de los nodos genera un impacto negativo en el desempeño y la estabilidad de una red P2P.
5. Generar un modelo que permita incentivar que los usuarios compartan sus recursos en escenarios de sobrecarga de la red tradicional.
6. Diseñar un algoritmo de bajo costo computacional que permita crear relaciones duraderas entre los nodos de una red P2P móvil, con el uso de un criterio previsor.
7. Comparar el efecto de los algoritmos de ruteo más usados en redes convencionales sobre redes P2P.
8. Brindar a la comunidad científica un aporte en relación con la búsqueda de una plataforma tecnológica que permita el acceso a las telecomunicaciones a poblaciones privadas de la infraestructura tradicional de telefonía.

1.3 Alcances y limitaciones

Esta investigación busca generar un modelo que permita analizar el comportamiento de las redes móviles y comparar los modelos de enrutamiento tradicionales con un algoritmo con criterio previsor. Son entregables dentro del alcance de esta investigación:

1.3. Alcances y limitaciones

1. Escenarios pequeños, medianos y grandes que sigan un patrón de movimiento convergencia/divergencia, escuadrón en movimiento y de concentración.
2. Un algoritmo para redes P2P con criterio previsor basado en tiempo y señal, tal como están definidas en este documento.
3. Un informe de análisis de complejidad temporal sobre el algoritmo.
4. El código fuente de un simulador funcional que reciba como entrada la descripción de la red y la evalúe al emplear los cuatro criterios (tiempo, señal, azar y previsor). Debe dar como resultado, en uno o varios archivos de texto plano:
 - El desempeño promedio de cada nodo en los diferentes pasos.
 - El desempeño instantáneo de la red en cada paso.
 - El desempeño final de la red en cada paso.
5. Artículo científico que se entregará al comité editorial de alguna revista o conferencia científica, con miras a su aprobación y publicación.
6. Análisis sobre el impacto del uso de criterio previsor en el desempeño y estabilidad de redes P2P móviles.

Esta investigación representa una fracción de todo el trabajo que se requeriría para implementar exitosamente una red P2P con criterio previsor para la selección de pares; sin embargo, por razones de tiempo, muchos aspectos no fueron tomados en cuenta para la presentación de tesis. La siguiente es una lista de las limitaciones:

- Consideraciones o la implementación del algoritmo en dispositivos móviles físicos.
- Consideraciones o la implementación del algoritmo para uso comercial.
- Patente alguna que prohíba el uso de cualquier aporte generado en este proyecto.

1.3. Alcances y limitaciones

- Consideraciones o implementaciones de aspectos de seguridad, como autenticación o cifrado de datos.
- Consideraciones acerca de la influencia del medio ambiente en la propagación y alcance de la señal.
- Cualquier otro resultado, documento, *software* o producción que no estén contemplados en los entregables.

En el capítulo 2 se brinda una definición de red P2P, se hace referencia al trabajo realizado en el campo y la importancia del presente aporte. El capítulo 3 se centra en el criterio previsor y su implementación. En el capítulo 4 se explican los experimentos con los que se puso a prueba la implementación. En el capítulo 5 se muestran los resultados obtenidos y su análisis estadístico. Finalmente, en el capítulo 6 se ofrece un espacio para la discusión, se exponen las conclusiones y el trabajo futuro.

Marco teórico y trabajo relacionado

“When the snows fall and the white
winds blow, the lone wolf dies but
the pack survives”.

George R.R. Martin.

El despliegue de antenas centrales de comunicación de telefonía móvil es bastante costoso en términos económicos; este se realiza al suponer que no todos los dispositivos móviles a su alcance van a gestionar una llamada al mismo tiempo, y el aprovisionamiento de canales se calcula con base en la cantidad esperada de llamadas concurrentes [7].

De acuerdo con los datos del informe sobre desarrollo humano 2010 de la Organización de las Naciones Unidas [15], un 96,9 % de la población costarricense tiene acceso a la telefonía, ya sea fija o móvil. Según el Banco Mundial [29], por cada 100 habitantes existen 146 conexiones de telefonía móvil en el país.

El incremento tan acelerado de los usuarios de servicios de telefonía móvil ha generado una gran dependencia de las actividades sociales y económicas del país en la plataforma que sustenta el servicio.

Eventos como catástrofes naturales no solo pueden dañar esta infraestructura, sino que generan un sobreuso de ella por parte de los usuarios que buscan conocer el estado de sus familiares, bienes y empresas. Si bien esta es una de las finalidades de la telefonía móvil, la sobrecarga de la infraestructura local puede afectar seriamente la disponibilidad del servicio para transferir llamadas de emergencia.

El 5 de setiembre del 2012, Costa Rica sufrió un terremoto de 7,6 grados que causó gran afectación en el área por la falta de suministro eléctrico, derrumbe de torres y sobredemanda del servicio [47]. Esta situación no ha sido el único reporte de no disponibilidad del servicio, pues lo mismo sucedió en el terremoto del 8 de enero del 2009 [30] (figura 2.1).

Otros eventos como las celebraciones de Navidad y Año nuevo también generan problemas a los operadores de telefonía móvil; por ejemplo, según datos del periódico *La Nación*, en el 2010 el Instituto Costarricense de Electricidad (ICE) empezó a alcanzar su tope de capacidad, con tres millones de líneas activas [44]. Si bien la capacidad de esta y otras empresas de la región ha ido creciendo, la demanda de servicios de voz y datos se ha incrementado en un número aún mayor.



Figura 2.1: Eventos como el terremoto de Cinchona, en Costa Rica, pueden desestabilizar las telecomunicaciones. Fuente: Wikimedia [5].

Por otro lado, en regiones del mundo donde la brecha tecnológica sigue siendo amplia, el acceso a las tecnologías de telecomunicación es poco [15] (figura 2.2). Dicho alejamiento obedece, en mayor parte, al alto costo de crear y operar una infraestructura de telefonía móvil tradicional, lo cual obliga a la comunidad científica a buscar nuevas formas descentralizadas para hacer accesibles estos servicios y abaratar los costos.

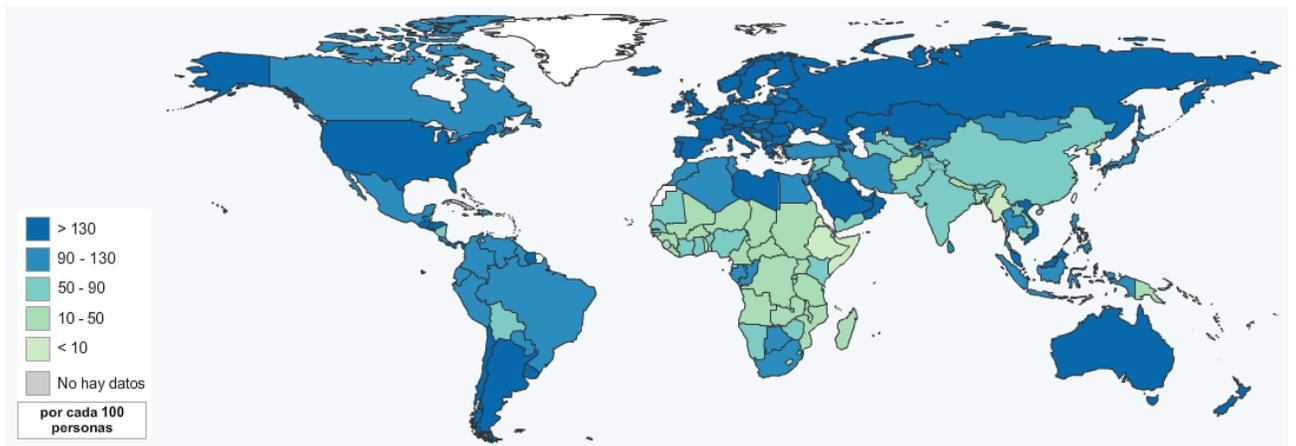


Figura 2.2: Abonados a telefonía fija o móvil por cada 100 personas. Fuente: Oficina del Informe sobre Desarrollo Humano [16].

De esta manera sería posible no solo brindar el servicio a países donde el costo del acceso a la infraestructura sea prohibitivo, sino permitir el despliegue de un sistema de telefonía móvil en zonas con una geografía que imposibilite la construcción de torres (figura 1.1, página 3).

En África, donde las redes fijas son insuficientes y en algunos casos inexistentes, las comunicaciones inalámbricas están creciendo de forma dramática y nuevos métodos de interconexión como las redes *mesh* se están volviendo más comunes [19].

2.1 ¿Qué es una red P2P?

Si bien el concepto de redes P2P tiene mucho tiempo de haberse ganado un espacio en el campo de la investigación, su comercialización y su uso para describir cualquier red que busca descentralización han hecho que no exista una definición universal de este [51]. La relación que se ha establecido entre el término P2P y la acción de compartir archivos por Internet hace difícil introducir el tema sin mencionar BitTorrent [59], Napster [60], Jabber [31] o Gnutella [26][22].

Esta investigación no pretende, de forma alguna, generar esa única definición, sino presentar una que sea consistente y pertinente con el campo de acción del presente trabajo. De esta forma, se busca apartar la confusión que generan otras definiciones, las cuales agregan elementos de otros modelos y generan un concepto híbrido o distraen al lector, pues solo enfocan la atención en el uso de estas redes (la definición de una red P2P debe estar separada de sus posibles usos, como compartir archivos o transmitir multimedios bajo demanda).

Por lo tanto, una red P2P se define como aquella conformada por elementos (nodos) cuya naturaleza física sea indiferente, pero su impacto sea potencialmente igual; es decir, cualquier nodo de la red debe ofrecer los mismos servicios que todos sus pares. Conceptos como supernodos, nodos administradores o repetidores no calzan con el

modelo que aquí se define.

Así pues, en una red P2P:

- Todos los nodos tienen los mismos permisos y responsabilidades [14].
- Su diseño asegura que cada usuario aporte recursos al sistema [14].
- Todos los nodos pueden solicitar y brindar el servicio a la vez [50].
- Cada nodo puede conectarse a la red a través de uno más nodos [61].
- La conexión entre dos nodos no depende de ningún nodo arbitrario [32] [50].
- Se asume que la infraestructura va a fallar en algún momento [61].
- Cualquier nodo escogido arbitrariamente puede ser removido de la red sin implicar ninguna pérdida en el servicio ofrecido por la red [50].

2.1.1 Teorema de CAP

De acuerdo con el modelo de Brewer (conocido como teorema de CAP) explicado en [24], una red distribuida solo puede asegurar totalmente dos de las siguientes características:

1. **Consistencia:** todos sus nodos visualizan los mismos valores para cualquiera de los datos disponibles en un tiempo determinado.
2. **Disponibilidad:** asegura que cada solicitud recibirá una respuesta y que toda operación sea finita.
3. **Tolerancia a fallos:** un sistema es tolerante a fallos si su operación continúa a pesar de que cualquiera de sus componentes deje de operar.

Consistencia

Es importante adelantar que la presente investigación no busca garantizar dicha característica; por lo tanto, no profundizará en ella. Sin embargo, es necesario hacer una breve explicación del tema para justificar por qué se decidió no garantizarla.

Existen varios modelos para asegurar consistencia en un sistema distribuido [1], tales como:

- Consistencia secuencial (estricta): el resultado de la ejecución de las operaciones de todos los nodos debe ser el mismo que si hubieran sido ejecutadas por uno solo [33].
- Consistencia débil: las secciones críticas del sistema son vistas por todos los nodos en un mismo orden secuencial. A su vez, todas las secciones críticas son vistas en el mismo orden secuencial en todos los nodos [18].
- Modelos relajados: existen muchos modelos que permiten un ordenamiento de las operaciones con una menor secuencialidad que los anteriores.

Si bien la consistencia es deseable en cualquier red, esta implica tiempo para sincronizar los nodos y su visualización de los datos. Aunque un algoritmo con criterio previsor busca el máximo beneficio local, no es un requisito que todos los nodos compartan una misma visualización con respecto al estado de la red; de hecho, no es necesario que un nodo tenga información acerca de los demás en la red, sino solo de aquellos que se encuentren (o hayan estado) dentro de su rango de alcance.

Además, como la conversación entre dos nodos es importante solo para estos, no existe una sección crítica de los datos, pues ningún tercero debería acceder a ella.

Disponibilidad

Para que el servicio que provee la red sea considerado disponible, cada una de las solicitudes que llegan a un nodo debe resultar en una respuesta [24]. Esto quiere decir

que cada operación en la red debe ser finita; sin embargo, no existe una limitación en cuanto al tiempo que pueda necesitar para terminar. Por esta razón, no es posible garantizar consistencia, pues para ello sí se requiere definir un máximo de tiempo.

Es importante notar que un mensaje de fallo es válido siempre y cuando sea generado por otro nodo. La disponibilidad es una de las características que se ha seleccionado garantizar durante esta investigación y, por lo tanto, es necesario establecer un mecanismo para evitar operaciones infinitas como la búsqueda de pares.

Tolerancia a fallos

Cuando un nodo falla, sus nodos vecinos deben ser capaces de establecer rutas a la red mediante el uso de otros nodos; a este proceso se le conoce como mecanismo de recuperación ante errores. Este es un paso obligatorio en cualquier red P2P y se han realizado muchos trabajos en la investigación de mejores técnicas para lograr dicho fin [61].

Es imposible eliminar la cantidad de fallos que un dispositivo va a presentar por un mal funcionamiento de *hardware* [13] [28]. Una de las principales razones por las que las redes P2P han cobrado tanta atención es que, a pesar del fallo de sus componentes, puede seguir operando y brindar el servicio, pues no depende de ninguno de estos nodos [52].

En el caso de redes P2P móviles, cuando un par que está siendo usado por un nodo para acceder a la red sale de su rango de alcance, se considera que ha habido un fallo, pues no hay forma de que el nodo sepa la causa por la que su par le dejó de brindar el servicio [48]. Así pues, el comportamiento mostrado en la figura 2.4 se considera un fallo del nodo superior por parte del nodo central.

Sin embargo, el tiempo que tarda un mecanismo de recuperación ante errores para encontrar una nueva ruta ante un fallo puede ser considerablemente largo, lo cual no es un comportamiento deseado en una red P2P para telefonía. En [61] el autor propone

un algoritmo para prevenir los fallos mediante el agrupamiento de nodos como medio de acceso a la red.

No obstante, agrupar nodos no asegura que los grupos no se alejen entre sí. Es preferible buscar una forma de minimizar la cantidad de fallos por alcance.

Una red que asegure la disponibilidad y tolerancia a fallos no solo es deseable, sino imprescindible, pues estos obstáculos son la constante del escenario por resolver y la razón misma que impulsa la búsqueda de una solución en este trabajo.

2.1.2 Despliegue de una red P2P

El proceso de despliegue (*bootstrapping*) busca que cada nodo descubra al menos otro nodo para que le ayude a ingresar al sistema; la idea es que el nodo par escogido sea de la mayor utilidad posible. Generalmente se consideran beneficiosos los nodos más cercanos; sin embargo, más adelante se verá que el beneficio en redes móviles es más difícil de medir.

De acuerdo con [12], existen varias formas de diseñar y ejecutar el despliegue para una red P2P:

- **Valor fijo:** cuando un nodo ingresa a la red, de antemano sabe con cuáles pares debe tratar para buscar su destino. Este esquema no resulta muy conveniente para una red donde los dispositivos están en constante movimiento y se desconoce cuáles van a ser los nodos más beneficiosos cuando ingrese a ella.
- **Re-uso:** cada nodo utiliza los últimos valores conocidos de sesiones previas en la red. Este esquema asume que el nodo ya ha ingresado previamente a la red, lo cual puede generar muchos errores.
- **Anuncio:** cuando un nodo ingresa a la red genera un mensaje *broadcast* y toma como pares iniciales a los primeros en contestar. La cantidad de pares por escoger es arbitraria y el tiempo de respuesta asume que estos son cercanos. Normalmente

se evita emplear dicho esquema, pues generar estos mensajes puede saturar de ruido la red. No obstante, cuando se hace referencia a dispositivos móviles, todo su funcionamiento se basa en mensajes de *broadcast*; por lo tanto, es algo normal en esta tecnología.

- **Servidor de anuncio:** este esquema resulta al combinar los anteriores, pues el nodo se conecta primeramente a un servidor ya conocido, y aparte de entregar su información, solicita la de sus pares. Obviamente, este método viola los principios de P2P aquí definidos [22].

2.1.3 Enrutamiento en una red P2P

Una vez que el nodo ingresó a la red, es necesario que utilice los recursos ofrecidos por sus pares para lograr su objetivo. La decisión sobre cuáles pares usar y en qué momento emplearlos es tarea de los algoritmos de enrutamiento. En redes tradicionales (TCP/IP), dado que no hay despliegue, se utilizan valores fijos o previstos por el usuario.

En redes P2P este no es el caso, pues los algoritmos de enrutamiento y sus valores pueden ser parte de los datos que un nodo adquiera. Como en las redes móviles los mensajes de *broadcast* son una constante, estos se podrían aprovechar para obtener nuevos datos acerca de los nodos vecinos y sus capacidades (en el capítulo 4 se detallan los criterios de selección por usar en esta investigación).

2.2 Tecnologías móviles e inalámbricas

Si bien en la presente investigación el alcance de esta etapa no incluye la implementación en un ambiente real (ver Alcances y limitaciones), es muy importante hacer una clara diferencia entre dispositivos inalámbricos y móviles [53] [49].

Un dispositivo inalámbrico es aquel que no necesita de una conexión cableada para establecer comunicación con otro dispositivo remoto; sin embargo, esto no significa que

el dispositivo se está moviendo mientras realiza dicha conexión ni que está diseñado con tales características. Por ejemplo, una computadora portátil está hecha para ser transportada a diferentes lugares, y una vez colocada en una superficie de trabajo puede conectarse inalámbricamente, pero su uso en movimiento no es recomendado, pues normalmente está fabricada con partes mecánicas que pueden causar daños al *hardware*.

Un dispositivo móvil es aquel que está hecho para ser utilizado y establecer conexiones mientras se desplaza de un lugar a otro; un teléfono móvil es un gran ejemplo de tal categoría. Así, es posible conceptualizar las plataformas móviles como un caso particular de las tecnologías inalámbricas.

Aunque el desarrollo de dispositivos y tecnologías móviles se ha vuelto tan popular que a veces cuesta distinguir la diferencia, esta aclaración es necesaria para hacer énfasis en que los dispositivos con los que eventualmente se pretende trabajar están en constante movimiento, por lo cual en la investigación debe reflejarse consciencia de este hecho.

2.3 Algoritmos con criterio previsor

Los algoritmos con criterio previsor (*foresighted*) se definen como aquellos que buscan el mayor beneficio a futuro en cada una de sus decisiones. En [34] los autores, mediante procesos de decisión de Markov, buscan maximizar la calidad del video a largo plazo, con lo que también se promueve compartir en una red P2P, lo cual es muy deseable [21] [27]. Así pues, descargar/compartir los segmentos de un video en orden secuencial no necesariamente reporta el mayor beneficio, pues muchas de las redes P2P propician compartir, mediante la asignación de mayor ancho de banda, a los segmentos con menor existencia en una red.

Este concepto puede ser portado a la selección de pares, donde se busca que cada relación entre dos pares sea duradera y de alto beneficio. Por ejemplo, en [22] los autores

proponen la selección de pares basada en el gusto musical. Al usar el campo designado para el género musical en los archivos con formato MP3 se pueden buscar pares cuyo contenido sea afín al gusto musical, en espera de que tengan el contenido que el nodo vaya a solicitar.

En [56] los autores proponen ErdOS, un sistema operativo que busca minimizar el uso de batería al compartir recursos entre los dispositivos cercanos con protocolos de bajo consumo como *bluetooth*. El algoritmo se describe como "distribución de los recursos móviles basada en oportunidad" (*opportunistic sharing of mobile phone resources*) y tiene en cuenta los siguientes retos:

- Descubrimiento adaptativo de los recursos: la búsqueda de recursos gasta batería. La información histórica y de contexto desempeña un rol muy importante.
- Selección del nodo correcto.
- Colaboración y agregado eficiente de recursos: la distribución eficiente de los recursos es vital para el buen funcionamiento de la red.
- Mecanismos IPC (Comunicación entre procesos, por sus siglas en inglés): usar el método de comunicación equivocado puede generar alta carga en el uso eficiente de la batería.
- Movilidad: los nodos pueden entrar y salir de la red, lo cual puede generar fallos en la experiencia del usuario final. Los autores proponen buscar un modelo predictivo de disponibilidad.

En [9] se busca generar este tipo de relaciones duraderas con un criterio previsor de disponibilidad. Los autores lograron encontrar que en una red como *eDonkey* es posible minimizar la cantidad de fallos, al seleccionar pares con un criterio previsor que tome en cuenta el comportamiento de un subconjunto de los nodos durante los últimos siete días.

Si bien los resultados de este proyecto son alentadores, siete días es un tiempo demasiado largo para empezar a prever el comportamiento de un nodo, en especial cuando ese lapso es superior al tiempo de vida esperado para toda la red que se describe en esta investigación.

Además, en el caso de redes móviles esta tarea no es trivial, pues el patrón de desplazamiento de dos dispositivos no está dictado por ningún algoritmo, sino por la voluntad de su dueño.

Así las cosas, es necesario buscar una forma de utilizar un criterio previsor que les permita a los nodos empezar a operar aun cuando solo se cuente con los datos de unas cuantas unidades de tiempo (ver definición de Tiempo). Por lo tanto, se debe definir cuáles datos usar y cómo medir el tiempo para saber si un criterio previsor puede lograr ese objetivo.

2.3.1 Criterios de selección

Se han seleccionado el tiempo y la intensidad de la señal como los criterios por comparar, pues estos datos son relativamente fáciles de conseguir en los dispositivos móviles actuales y son muy usados tanto en redes convencionales como en redes P2P.

Señal

Se define la señal del dispositivo como la intensidad con la que su antena puede irradiar y establecer una comunicación sin generar pérdida. En redes no móviles, el ancho de banda (medido en Mhz) y la capacidad del canal (medida en Mbps) son los factores que normalmente toman mayor peso en la decisión de una ruta.

Aunque estas medidas también existen y son usadas en las redes móviles, hay mayor facilidad para asignar ancho de banda (por ejemplo, la banda ISM), y la capacidad máxima del canal depende en gran medida del *hardware* que se utilice, por lo cual su estudio se deja de lado por el momento. De esta manera, se analizará cada nodo de

la red como una antena isotrónica [36] para eliminar el efecto que pueda generar la variación en el patrón único de un irradiador.

Si los resultados de esta investigación generan evidencia estadística necesaria para considerar que los efectos del criterio previsor son superiores a los criterios convencionales, el estudio de este con patrones específicos de antenas móviles y las variaciones generadas por los modelos de dispositivos en el mercado serían los pasos por seguir en la investigación (ver Alcances y limitaciones).

Tiempo

Se define tiempo como la forma en que se mide el estado de la red; así, es posible detener su ejecución, generar una imagen de su estado y compararla con imágenes anteriores. Cada una de ellas estará separada de sus predecesoras por el mismo valor de tiempo y será llamada **paso**.

Esta decisión no solo permite un mejor modelado e interpretación de resultados, sino un mejor análisis de complejidad temporal. El tiempo es otra medida muy común en el desempeño de redes convencionales y se establece en otras medidas como latencia y saltos (*hops*).

Si bien estas medidas han sido usadas desde hace varias décadas en los algoritmos de enrutamiento [54, p. 382], las redes móviles presentan retos que hacen a la señal y al tiempo insuficientes como únicos criterios para seleccionar el mejor camino para elegir. La señal ya no será constante en el tiempo, pues en realidad el rango de señal que cada nodo ofrezca ya no dependerá solo de sus características de *hardware*, sino de la ubicación de este en un punto arbitrario del tiempo.

En la figura 2.3 se indica cómo un nodo podría tomar la decisión menos adecuada al elegir al nodo (superior) con mayor señal en un paso t e ignorar que este tiene poco tiempo de haber entrado a su rango de alcance. Aunque en el paso $t + 1$ la escogencia sigue pareciendo acertada, el paso $t + 2$ muestra lo contrario, pues el nodo seleccionado

sale permanentemente del alcance del nodo central.

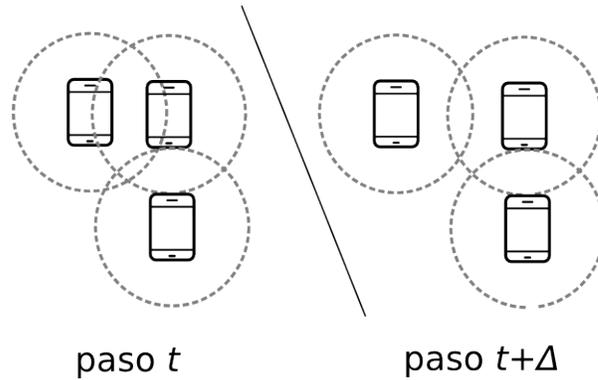


Figura 2.3: La intensidad de la señal de un dispositivo móvil depende de la ubicación de este en un punto determinado del tiempo.

En la figura 2.4 se muestra que analizar solo la fuerza de la señal puede ser un grave error en redes móviles, pues el nodo central escogería al nodo superior sobre el inferior en el tiempo t , por presentar una señal más intensa. Sin embargo, que la señal sea fuerte en un momento del tiempo no asegura que se mantendrá en el futuro. Cuando el nodo superior cambia de posición en $t + 2$, fuerza al central a elegir un nuevo camino para el envío de los datos, y probablemente sea el inferior. Este proceso de cambio de ruta puede generar interrupciones en la comunicación, sin mencionar renegociación de sesiones en capas superiores de *software*.

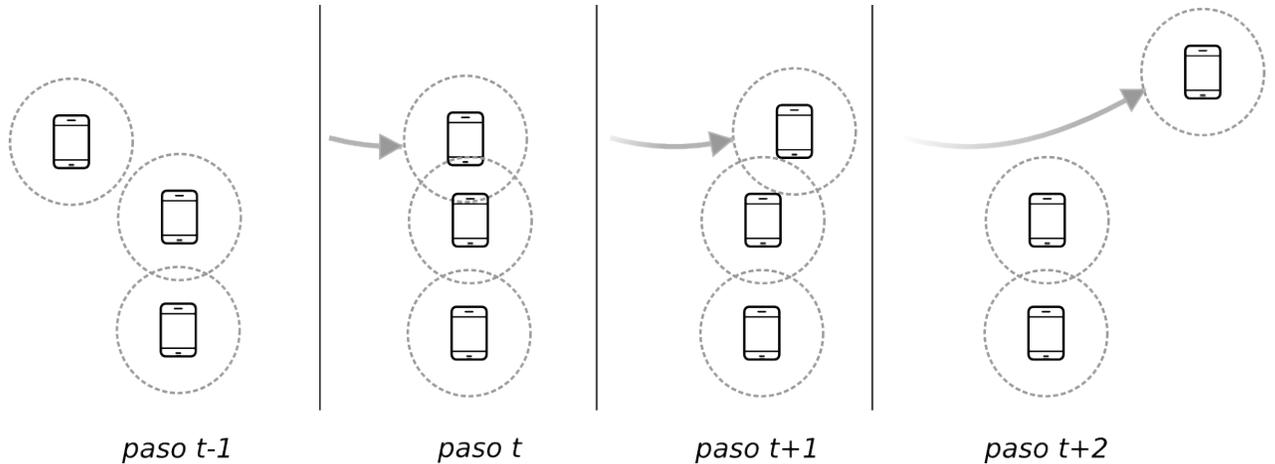


Figura 2.4: Seleccionar a un nodo solo por tener una señal más fuerte no necesariamente es el mejor criterio, pues esta podría ser poco duradera.

El caso podría empeorar con la aparición de nodos satélite que repetidamente se acercan al radio de alcance tan solo para desaparecer de este en un tiempo muy corto. Como se consideró previamente en el apartado *Tolerancia a fallos*, este comportamiento implicaría que el mecanismo de recuperación a fallos se ponga en funcionamiento repetidas veces, lo cual minimiza el desempeño de la red.

Así mismo, seleccionar como ruta principal aquel nodo que haya establecido la relación más larga no asegura que sea el mejor candidato. Nuevos nodos podrían estar ingresando a la red en el radio de alcance y presentar mejor señal durante un tiempo suficiente para creer que se mantendrá dentro del alcance en el futuro.

2.4 Otras propuestas en el campo

El tema de redes P2P no es nuevo en el ámbito de la investigación en telefonía. Existen muchas propuestas, algunas de ellas han salido del campo teórico para generar un impacto en el campo socioeconómico.

Compañías como Fon®[20] y Ooma®[40] ofrecen soluciones de telefonía mediante redes P2P que a primera vista parecieran ser móviles, pero realmente no lo son. Los

teléfonos se conectan a una antena fija en un dispositivo con conexión a Internet, y mediante una red de pares alcanza a su destino.

Si bien estas aplicaciones son funcionales, no son tolerantes a fallos, pues el dispositivo al que se conectan los teléfonos presenta las mismas limitaciones de una red convencional de telefonía móvil (figura 2.5), donde se limita el punto de acceso a la red a un único nodo que, en caso de fallo, separaría a los dispositivos no solo de la red, sino entre ellos.

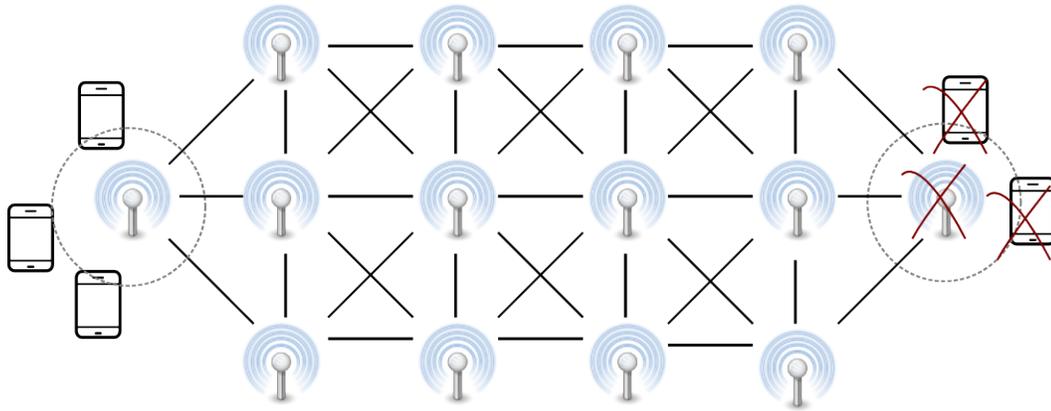


Figura 2.5: En un modelo híbrido, el fallo de una estación implica la desconexión total de los dispositivos. Estos no solo quedan aislados de la red, sino también imposibilitados de comunicación alguna entre ellos.

Además, estos productos minimizan el rango de movilidad de los dispositivos (figura 2.6). Si un dispositivo se aleja de la estación para la cual está configurado utilizar como puerta de acceso, queda totalmente aislado de la red, aun cuando se encuentre al alcance de otro dispositivo o estación que sí tenga acceso a la red.

Por lo tanto, en el último tramo del esquema, los dispositivos desempeñan un rol secundario que los convierte en clientes de la red y no en proveedores, lo cual va en contra del espíritu de una red P2P que siga la definición presentada aquí.

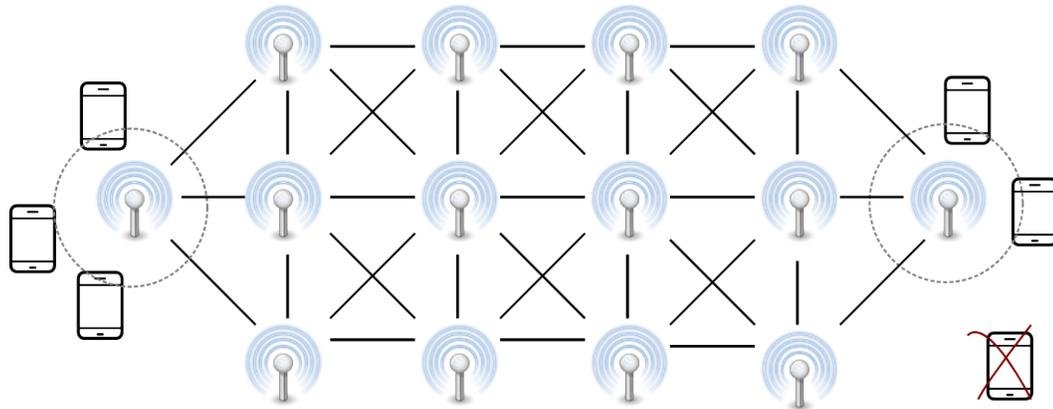


Figura 2.6: En un modelo híbrido, la movilidad de un dispositivo se ve reducida por el alcance de su estación.

Otras propuestas como OpenVoip [42] han intentado delegar la solución a otras capas y protocolos como P2PP y P2PSIP [12], cuya solución para un fallo en la red es que el nodo salga y reingrese por otra parte de esta. Lamentablemente, el proyecto cayó en el abandono, a pesar de los grandes aportes que hizo al campo.

Por su parte, Jabber [31] es un ejemplo muy citado de redes P2P, no solo por la gran aceptación que ha tenido en el mercado, sino por ser un protocolo libre. Si bien el proyecto permite la transmisión de audio en tiempo real, no fue diseñado para dispositivos móviles y necesita de un despliegue con valor fijo (el cliente necesita saber a cuál servidor conectarse).

Finalmente, es importante el mencionar el proyecto de TerraNet®[55], que tiene un objetivo muy alineado con el de la presente propuesta. Esta empresa, junto con el apoyo de otras entidades como Intel®, propone crear una plataforma tecnológica que les permita a dispositivos ya existentes (teléfonos con sistema operativo Linux/Android) desplegar una red como la que se propone en este documento. Sin embargo, a la fecha no fue posible hallar ninguna publicación del proyecto que haga un aporte al conocimiento científico, por lo que a corto plazo no será posible comparar ambas iniciativas.

2.5 Importancia

El principal beneficio de la telefonía moderna es la movilidad. Poder trasladarse sin perder la capacidad de comunicarse con otras personas (que también podrían estarse moviendo) mediante el uso de algún dispositivo de *hardware* era poco común hace un par de décadas. La gran explosión que estas tecnologías tuvieron en el mercado permite que en la actualidad sea considerado asocial no tener un teléfono móvil en muchos países del mundo.

Sin embargo, esa demanda tan acelerada de servicios de telecomunicación está empezando a alcanzar los límites de la oferta que la infraestructura tradicional puede ofrecer. Por eso es importante buscar una forma escalable de ofrecer el servicio, la cual se adapte a los patrones de movilidad de las diferentes poblaciones humanas.

Si bien existen varias propuestas para el uso de redes P2P de telefonía, la mayoría de las que ofrecen soluciones listas para su uso asume un escenario no cambiante, en el que cada usuario tiene una conexión a Internet estable o se va a encontrar en un área determinada.

Este trabajo no pretende minimizar la importancia de los proyectos mencionados, pues han mostrado gran utilidad en algunas zonas rurales y alejadas de los medios de comunicación convencionales. No obstante, estas soluciones limitan a los usuarios a un área de movilidad muy pequeña y poco estable. Por lo tanto, el despliegue de una solución debe ser ágil, y es necesario asumir que las condiciones van a cambiar con rapidez, incluso durante el despliegue mismo.

Las otras propuestas analizadas que buscan tratar problemas similares a los de esta investigación, donde cada dispositivo es un individuo de primer mundo cuyo entorno es difícil y cambiante, han delegado estos problemas a otras capas de *software* que hasta el momento no han sido modificadas para encarar los retos de una red P2P para dispositivos móviles.

Así, esta investigación busca presentar una nueva propuesta que permita hacer un

2.5. Importancia

despliegue rápido de una red P2P con base en el uso de teléfonos y dispositivos móviles, sin la necesidad de una torre central de comunicaciones que funcione a pesar del fallo de cualquiera de sus nodos.

A nivel científico, los resultados del presente proyecto podrían ser exportados a otros campos como el de enrutamiento en redes convencionales, que actualmente está siendo dirigido más por principios económicos que científicos [2], con lo cual se deja de lado el desempeño y la estabilidad.

A nivel social, se apuesta fuertemente por un modelo que permita generar una herramienta vital en el trabajo de socorristas, fallos en la capacidad de la actual infraestructura de comunicaciones y poblaciones alejadas de los medios de comunicación tradicionales (figura 2.7).



Figura 2.7: Los escenarios más beneficiados por la investigación pueden ser: a) redes convencionales congestionadas durante desastres naturales o eventos sociales b) equipos de rescate c) poblaciones humanas alejadas de las redes móviles convencionales.

Criterio previsor: desarrollo y análisis

“Each problem that I solved became a rule, which served afterwards to solve other problems”.

René Descartes

Con base en la motivación e importancia del proyecto, así como en la definición de una red P2P ya presentada, se propone crear una red P2P móvil que, mediante el uso de un algoritmo de selección de pares basado en un criterio previsor, busque minimizar la cantidad de fallos por pérdida de señal entre los nodos y además maximice el desempeño de la red durante todo su tiempo de vida.

3.1 Simulador

Uno de los aportes de esta investigación es concluir si es viable la construcción de redes P2P como la aquí descrita. Una forma de hacerlo es al implementar la propuesta mediante una aplicación de *software* que pueda ser distribuida a la población de usuarios móviles y determinar su comportamiento; sin embargo, no hay certeza alguna del número de usuarios que sería posible conseguir, ni la aceptación de estos. Otra forma sería comprar un gran número de dispositivos móviles y utilizar la implementación antes descrita, pero el costo resultaría altamente prohibitivo e igualmente sería necesaria la participación de una gran cantidad de personas para que utilicen los dispositivos.

Además, en esta investigación aún no se han considerado muchos factores (ver Factores y niveles), por ejemplo, si la implementación debe ser parte del núcleo del sistema operativo, correr en el espacio de usuario o incluso ser parte del protocolo de comunicaciones (implementado en *hardware*). Asimismo, existen muchos modelos y fabricantes de dispositivos móviles, cada uno con diferentes características como el patrón de irradiación, probabilidad de fallo por *software*, probabilidad de fallo por *hardware*, entre otros.

Por lo tanto, antes de implementarlo en alguna plataforma móvil, y por limitaciones tanto de presupuesto como de tiempo, se ha decidido centrar esta investigación en generar un simulador que permita determinar estadísticamente la veracidad de la hipótesis, y por consiguiente, la viabilidad de utilizar tal propuesta en escenarios reales.

Este simulador es capaz de utilizar la descripción de cada población (tamaño y patrón de migración) como entrada para su ejecución, y mantener para cada nodo su selección de pares con base en el uso de los criterios de tiempo, señal, azar y previsor.

La siguiente es una lista de las principales características del equipo usado para el desarrollo del simulador y ejecución de las pruebas:

- Intel(R) Core(TM) i7-3770 CPU @ 3.40GHz (8 núcleos).
- 12 Gb de memoria Ram.
- 1 Tb de espacio en disco duro.
- Sistema operativo GNU/Linux: Debian Jessie X86_64.
- C y C++ como lenguaje de programación.
- Compilador gcc versión 4.8.1 (Debian 4.8.1-10).

Se desarrollaron dos versiones del simulador. En la primera iteración se utilizó PEDSIM [25]. El simulador es funcional y genera datos que pueden ser utilizados para esta investigación; sin embargo, si bien muchas de las rutinas internas de PEDSIM proveen las funcionalidades que se requerían para el proyecto (por ejemplo, calcular el conjunto de vecinos para un nodo arbitrario), no todos los datos necesarios eran parte del resultado (las distancias entre esos vecinos y el nodo). Esto hizo necesario recalcular muchos datos y repetir operaciones, lo cual dio como consecuencia un simulador que tardaba hasta tres semanas en ejecutar una ronda de simulaciones. La información correspondiente PEDSIM se encuentra en el anexo D.

La segunda iteración se realizó al emplear RVO2 [39] como base para el simulador. Con esta implementación fue posible obtener los resultados de las simulaciones en cuestión de horas.

Finalmente, fue necesario crear e incorporar herramientas para sintetizar y analizar los resultados después de cada ejecución del simulador. Como se esperaba una cantidad

bastante grande de datos en los resultados de las simulaciones, se crearon *scripts*¹ para procesar dicha información.

3.1.1 RVO2

Las simulaciones de múltiples agentes que comparten un espacio de trabajo común han ganado cada vez más atención para diversos fines, tales como simulación de multitudes, la navegación por un equipo de robots móviles, videojuegos, estudio de fenómenos naturales, ingeniería de tráfico, la arquitectura y el diseño, simulaciones de formación de emergencia, entre otros [39].

RVO2 es un algoritmo para la navegación interactiva y la planificación de un gran número de agentes en entornos bidimensionales (multitudes). En tiempo de ejecución, cada agente detecta el medio ambiente de forma independiente y calcula un movimiento libre de colisiones basado en la prevención de colisiones recíproca (ORCA, por sus siglas en inglés). El algoritmo asegura que cada agente no exhibe comportamientos oscilatorios.

La biblioteca para C++ es una aplicación de dicho algoritmo. Tiene una sencilla API para aplicaciones de terceros. El usuario especifica los obstáculos estáticos, agentes y las velocidades preferidas de los agentes. La simulación se realiza paso a paso a través de una simple llamada a la biblioteca; es totalmente accesible y manipulable durante la ejecución. La biblioteca explota varios procesadores si están disponibles y usa OpenMP para la paralelización eficiente de la simulación.

Al poder manipular tan fácilmente las rutinas de RVO2, se desarrolló un simulador que es capaz de extraer los datos pertinentes para el análisis posterior.

¹Aunque la traducción “archivo de ejecución por lotes” es correcta, el término *script* es de mayor aceptación; por lo tanto, será el que se utilice en este documento.

3.1.2 Escenarios

En el anexo B es posible encontrar un *script* para *bash* [23] que modela escenarios en archivos XML [57], los cuales después pueden ser consumidos por el simulador desarrollado.

Fue necesario desarrollar 3840 escenarios (capítulo 4) como los descritos en el capítulo 2 para ser probados con los criterios de tiempo, azar, señal y finalmente con el previsor. Los escenarios se han clasificado en tres categorías de tamaño:

- **Pequeños:** 0-100 nodos. Son útiles para desarrollar la prueba de concepto y afinar detalles en la implementación del algoritmo.
- **Medianos:** 100-500 nodos. Se desea conocer cuál es el mínimo de nodos necesarios para alcanzar un comportamiento estable.
- **Grandes:** 500-2000 nodos. Se desea conocer si una sobreexistencia de disponibilidad puede afectar negativamente el desempeño final del algoritmo.

3.1.3 Patrones de movimiento

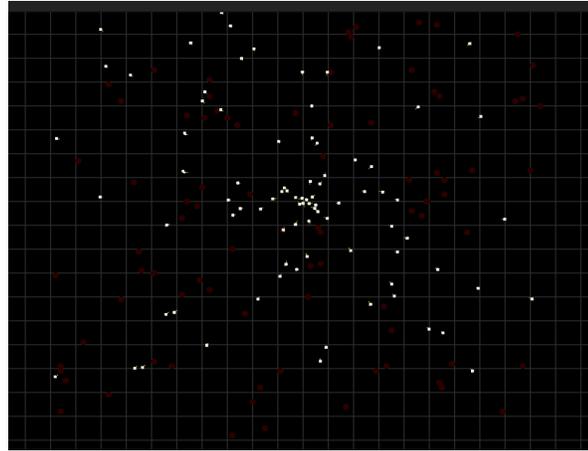
El patrón de movimiento de un nodo en la red es un tema complejo, especialmente cuando esta representa una población específica en la que cada nodo tiene autonomía sobre su desplazamiento. En este proyecto, los escenarios pueden representar a uno de tres tipos distintos de poblaciones, y se ha escogido un patrón de movimiento entre los propuestos en [17].

Convergencia/divergencia

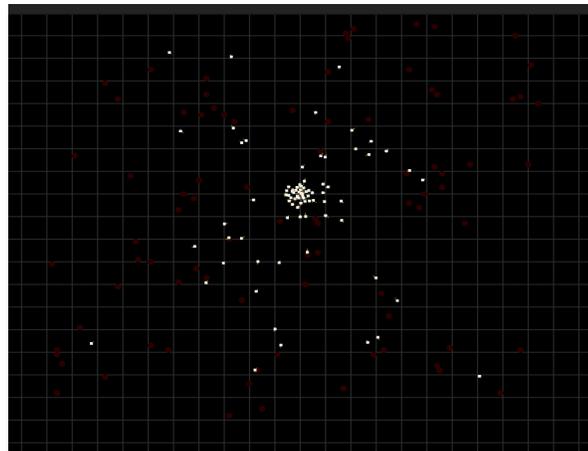
Es el caso más típico de ciudades y sectores donde la población está concentrada en un área relativamente pequeña. Los individuos también presentan un patrón repetitivo, en el cual las rutas que trazan convergen durante el día (figura 3.1b) a los centros de trabajo y estudio, pero al final del día divergen a sus hogares. Para este caso se ha seleccionado el patrón de convergencia/divergencia.

Para su generación se crea un punto de ruta común a todos los agentes. Cada uno de ellos tiene ubicación inicial aleatoria y un segundo punto de ruta. El agente se desplazará cíclicamente entre esos dos puntos de ruta. El anexo C muestra un ejemplo de dicho escenario.

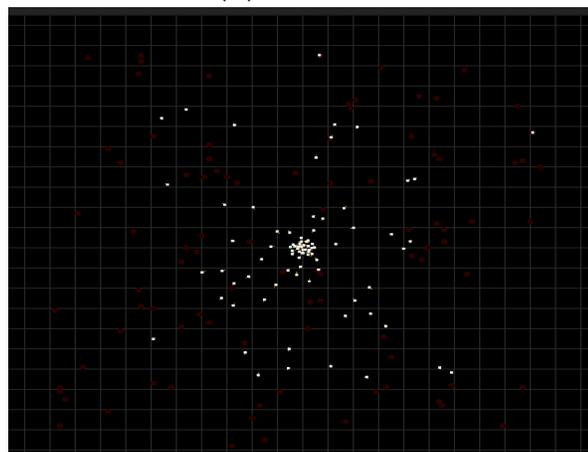
Es importante notar que el proceso es continuo y la aceleración de los agentes es variable, por lo cual los ciclos de estos se mantendrán activos y con tiempos distintos. Por eso, va a suceder que no todos se concentrarán y esparcirán al mismo ritmo, como se puede apreciar en la figura 3.1c. Esta simulación corresponde al tiempo en que una ciudad presenta mayor actividad.



(a) Paso t



(b) Paso $t+k$



(c) Paso $t+k+p$

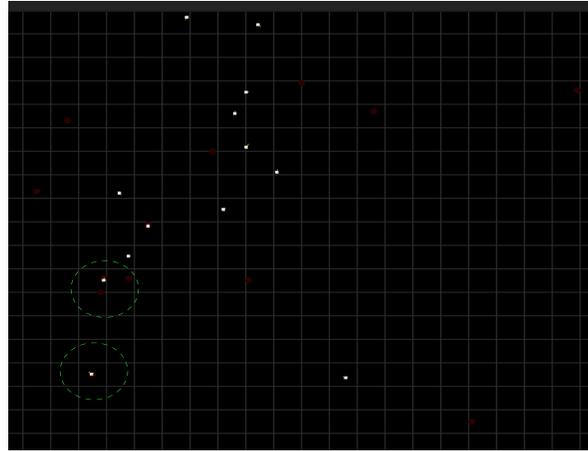
Figura 3.1: Ejemplo de la ejecución de un escenario con patrón de convergencia/divergencia en tres pasos distintos.

Concentración

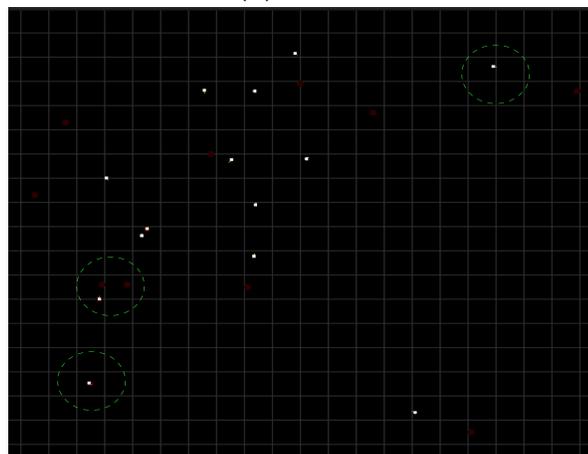
Por su alejamiento de los medios de comunicación y transporte convencionales, los habitantes de poblaciones rurales se mantienen en movimiento, tanto en el día como en la noche, pero dentro de un rango pequeño, posiblemente cercano a sus hogares.

Es semejante a un escenario de convergencia/divergencia, pues los agentes son modelados con una ubicación inicial aleatoria y dos puntos de ruta. Sin embargo, no se fuerza a que todos tengan un punto de ruta común. Esto genera el comportamiento deseado, donde los agentes se desplazan y siguen una rutina probablemente muy distinta a la de los demás agentes.

En la esquina inferior izquierda de las figuras 3.2a y 3.2b puede apreciarse un agente que solo tiene un punto de ruta y se mantiene estático. El que está en la posición superior tiene dos puntos de ruta y se mantiene entre ellos. No obstante, en la figura 3.2b se puede apreciar un tercer agente en la esquina superior derecha, que está cruzando el campo con una ruta más larga.



(a) Paso t



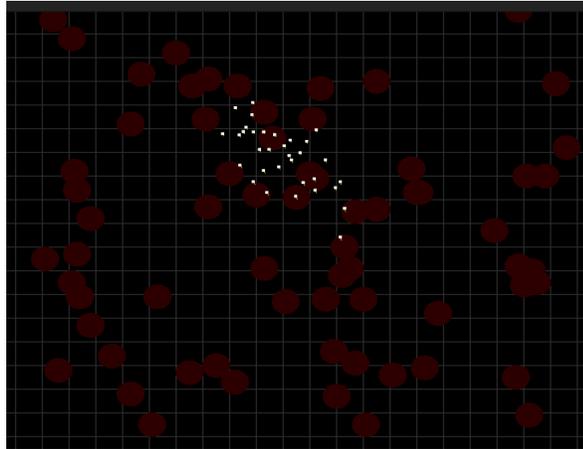
(b) Paso $t+k$

Figura 3.2: Ejemplo de la ejecución de un escenario con patrón de concentración en pasos distintos.

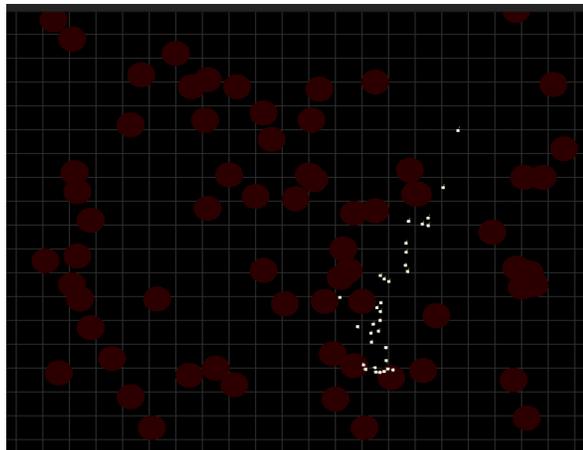
Escuadrón

Si bien el patrón de un escuadrón de rescate puede variar según el tipo de emergencia y la zona, por lo general estos equipos buscan barrer un área determinada en busca de sobrevivientes. Para este escenario se ha seleccionado el patrón de movimiento de escuadrón en movimiento (*moving cluster*).

En su modelado, todos los agentes comparten la misma ruta. Su ubicación inicial será cercana, pues todos serán ubicados a una distancia máxima aleatoria a partir del punto de ruta inicial.



(a) Paso t



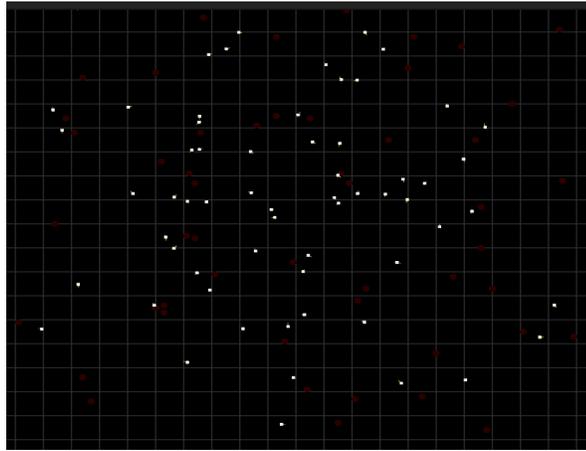
(b) Paso $t+k$

Figura 3.3: Ejemplo de la ejecución de un escenario con patrón de escuadrón en pasos distintos.

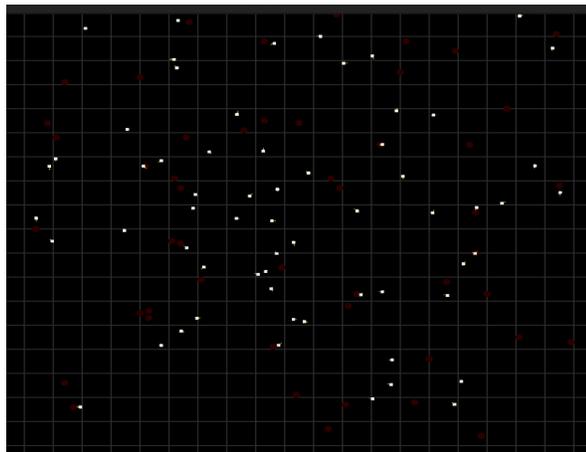
Aleatorios

A diferencia de los anteriores, el patrón aleatorio no se encuentra listado en [17]. Sin embargo, como buena práctica de DoE (capítulo 4) es conveniente incluir casos aleatorios para establecer una base de comparación con los efectos de los factores por analizar.

La cantidad de puntos de ruta por elegir, de agentes y los puntos de ruta que elija cada uno de ellos es generada de forma aleatoria. Al igual que en los otros escenarios, la ubicación inicial de cada agente es aleatoria. En la figura 3.4 se pueden apreciar dos capturas en tiempos distintos para un escenario aleatorio.



(a) Paso t



(b) Paso $t+k$

Figura 3.4: Ejemplo de la ejecución de un escenario con patrón aleatorio en pasos distintos.

3.2 Desarrollo del criterio previsor

Dado que el tiempo y señal son los datos más básicos de un dispositivo móvil, utilizar cualquier otra fuente de información supondría una evidente ventaja para el algoritmo y afectaría su credibilidad. Por lo tanto, el algoritmo debe estar en función de estos dos parámetros, donde cada nodo analiza a sus vecinos (los nodos en su rango de alcance) y les asigna una puntuación.

Para el desarrollo de las siguientes secciones es necesario establecer unas definiciones previas:

Definiciones 1. *Nodos*

- *Sea R la red.*
- *$A_i = \langle X_i, Y_i \rangle$, donde $\langle X_i, Y_i \rangle$ son las coordenadas del nodo A_i .*
- *$A_i \in R$.*
- *$|R| = n$.*

Definiciones 2. *Métricas*

- *Sea V_i el vector con los tiempos para $A_k \in R$ tal que $A_i \neq A_k$.*
- *$|V_i| = n - 1$.*
- *Sea F_i el vector con los puntajes asignados por A_i para $A_k \in R$ tal que $A_i \neq A_k$.*
- *$|F_i| = n - 1$.*

Definiciones 3. *Vecinos y pares*

- *Sea $E(A_i, A_k)$ la distancia euclidiana [8] para las coordenadas de los nodos A_i, A_k .*
- *Sea d_i la distancia a la que la señal de A_i puede funcionar.*

3.2. Desarrollo del criterio previsor

- $\forall A_k \in R$ tal que $A_k \neq A_i$, $e(A_k, A_i) < d_i \Rightarrow A_k$ es vecino A_i .
- Sea $A_p \in R$ el nodo par escogido por A_i .
- $A_p \neq A_i$.
- A_p es vecino de A_i .

Este criterio previsor se formuló como un algoritmo cuyo consumo de recursos computacionales pueda ser analizado y comparado al utilizar la teoría y técnicas de análisis de algoritmos [3].

Las siguientes fórmulas muestran el algoritmo desarrollado:

$$peer(A_i) = \begin{cases} A_p, & \text{si } A_p \neq \emptyset \wedge e(A_p, A_i) \leq d_i \\ peerfs(A_i), & \text{de lo contrario} \end{cases}$$

$peerfs(A_i)$:

$$\forall k = 1, 2, \dots, |R| \text{ tq } k \neq i \quad \begin{cases} V_i[k] + 1 \wedge F_i[k] \leftarrow score(A_i, A_k), & \text{si } A_k \text{ es vecino de } A_i \\ V_i[k] \leftarrow 0 \wedge F_i[k] \leftarrow -1, & \text{de lo contrario} \end{cases}$$

$$A_p \leftarrow \max(F_i[1], \dots, F_i[|R|])$$

$$\rightarrow A_p$$

$$score(A_i, A_k) = 0.5 \cdot e^{-4 \cdot e^{-0.5V_i[k]}} + 0.5 \cdot e^{-0.1 \cdot e^{0.5E(A_k, A_i)}}$$

3.2.1 Función de puntuación

La función de puntuación es la pieza clave para el funcionamiento del algoritmo, pues es la que inyecta el criterio previsor.

$$score(A_i, A_k) =$$

$$0.5 \cdot e^{-4 \cdot e^{-0.5V_i[k]}} + 0.5 \cdot e^{-0.1 \cdot e^{0.5E(A_k, A_i)}}$$

Está basada en la función de Gompertz [6], que fue creada inicialmente para predecir la mortalidad. Esta función sigmoide crece poco al inicio y final de su curva, tal como se puede apreciar en el ejemplo de la figura 3.5.

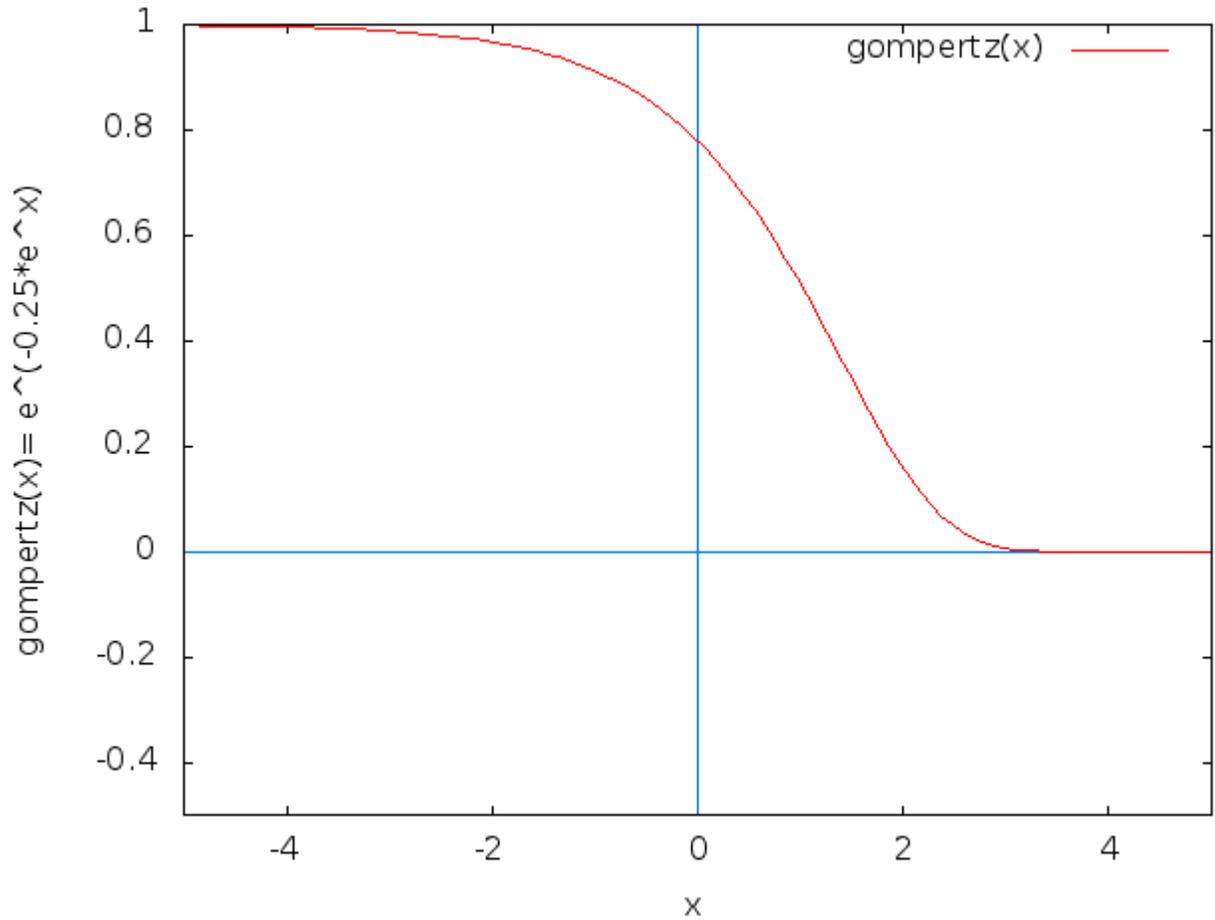


Figura 3.5: Gompertz es una función sigmoide.

Ese comportamiento es deseable, pues al inicio no es posible saber con certeza que el nodo vecino ha estado lo suficiente como para seguir siendo el mejor candidato. Después de muchos pasos ya se ha comprobado que es un buen candidato, por lo que no es necesario agregarle mucha puntuación durante los pasos siguientes.

A la hora de puntuar la señal se trata con el caso contrario, pues se definió la distancia euclidiana como la métrica para calcular su alcance. A mayor distancia, menor fuerza de señal. En la figura 3.6 se puede apreciar el comportamiento de ambas sigmoides, cuyo puntaje puede ser de $[0, 0.5]$. Al combinar estas funciones se obtiene el comportamiento de la figura 3.7, cuyo dominio comprende el intervalo $[0,1]$.

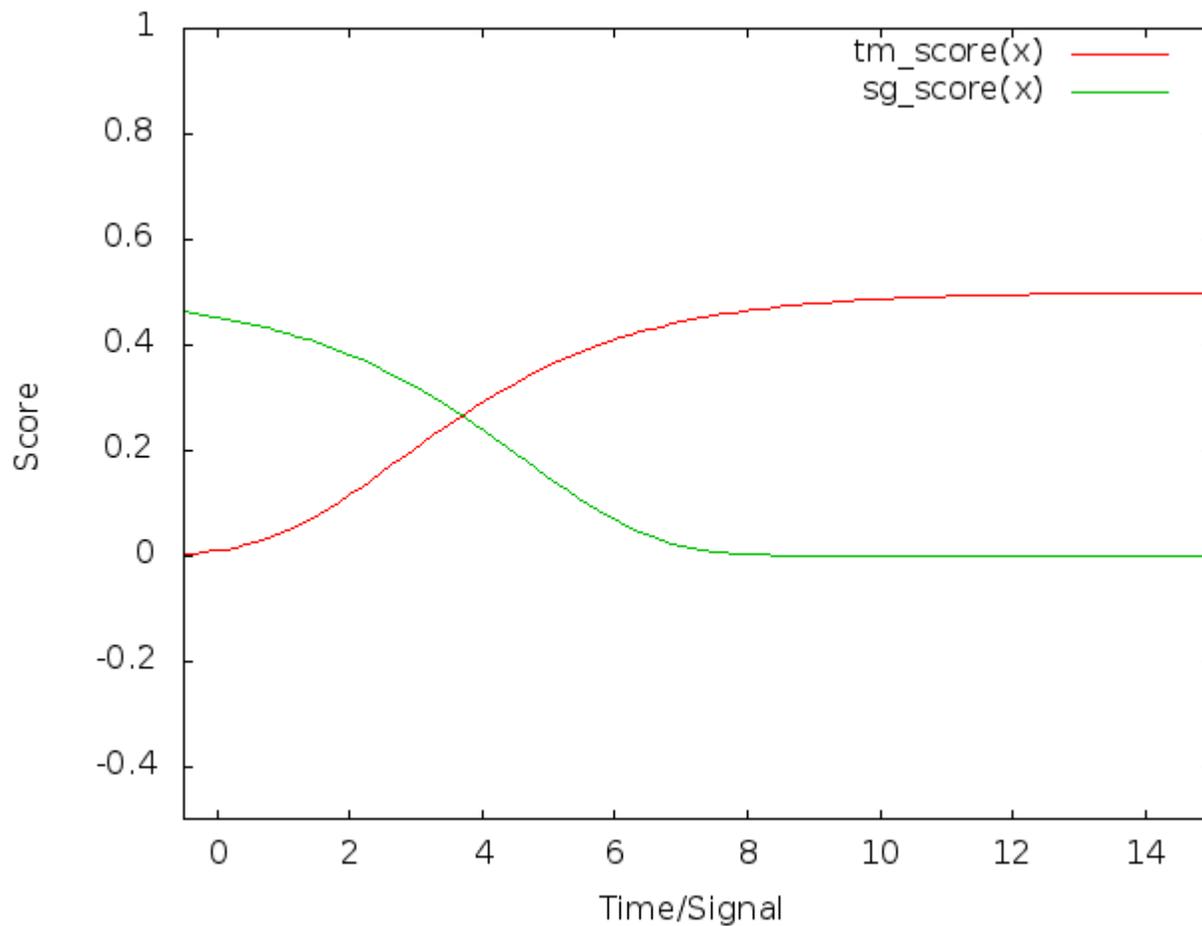


Figura 3.6: Proyección de las funciones de puntaje para el tiempo (tm_score) y para la señal (sg_score) usadas en la función global de puntaje previsor. X representa la intensidad del parámetro de entrada, mientras que Y representa el puntaje.

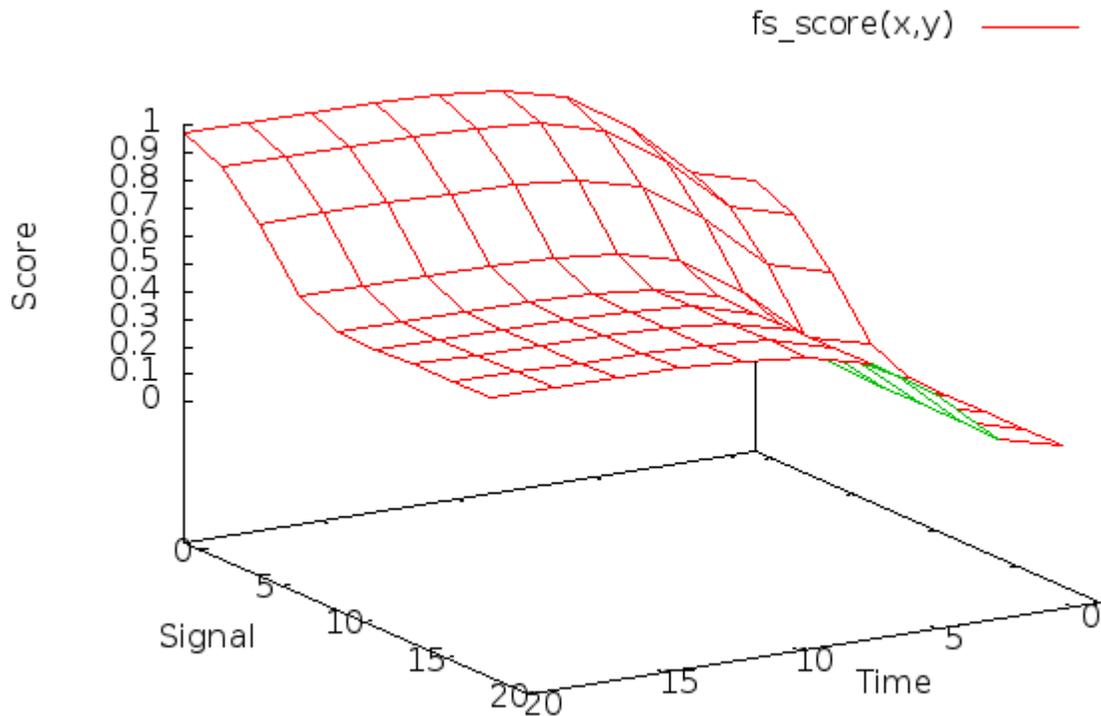


Figura 3.7: Proyección de la función de puntaje predictor. Su resultado varía de acuerdo con los parámetros de tiempo (time) y señal (signal).

3.3 Análisis de complejidad

El uso eficiente de los ciclos de procesador y la memoria es de gran importancia para esta investigación. Independientemente de los resultados, es necesario garantizar que el uso de recursos sea bajo, pues se desea implementar en dispositivos con características limitadas (CPU y memoria).

Para las siguientes secciones se seguirán usando las definiciones 1 de la sección anterior. Además, se acotará la complejidad a sus niveles superiores [3].

3.3.1 Análisis de complejidad espacial

Los tamaños de las estructuras de datos A_i , V_i y F_i no son modificadas por el algoritmo, lo cual reduce significativamente su consumo de memoria y facilita su análisis. Incluso al ejecutarse en una red distribuida, cada nodo A_i utiliza únicamente los vectores V_i y F_i correspondientes, y así se minimiza el impacto:

$$\begin{aligned}\forall A_i \in R \quad \exists \quad V_i, F_i \\ |A_i| &= 2 && \text{(por 2)} \\ |F_i| &= n - 1 && \text{(por 2)} \\ |V_i| &= n - 1 && \text{(por 2)} \\ \therefore \Theta(\text{peers}) &= n - 1 + n - 1 + n * 2 \\ &= 4n - 2\end{aligned}$$

Entonces se necesitan $4n - 2$ unidades de espacio por nodo. Así, por ejemplo, en una arquitectura con enteros de 64 bits y una red de 5000 nodos se ocuparían $((4 * 5000 - 2) * 64) \text{bits} = 1.279872e6 \text{bits} = 159984 \text{bytes} = 156 \text{Kb}$, en caso de que el dispositivo esté al alcance de todos los demás nodos de la red.

Al utilizar asignación de memoria dinámica, fácilmente se puede reducir el uso de memoria en cada dispositivo, para mantener en memoria solo los datos de aquellos nodos que estén a su alcance.

3.3.2 Análisis de complejidad temporal

El buen uso de CPU en los dispositivos móviles es muy deseado por una comunidad que exige cada vez más tiempo de vida para su batería [45] [19] [56]; de ahí que esta investigación busque un algoritmo de bajo costo computacional.

$$peer(A_i) = \underbrace{\begin{cases} A_p, & \text{si } A_p \neq \emptyset \wedge e(A_p, A_i) \leq d_i \\ peerfs(A_i), & \text{de lo contrario} \end{cases}}_{\Theta(k) + \Theta(peerfs(A_i))}$$

$$\begin{aligned} \therefore \Theta(peer(A_i)) &= \Theta(k) + \Theta(peerfs(A_i)) \\ &=? \end{aligned} \tag{3.1}$$

La primera decisión del algoritmo tiene un tiempo constante; sin embargo, la segunda no. Entonces es necesario analizar la complejidad temporal para la segunda decisión posible.

$peerfs(A_i)$:

$$\underbrace{\forall k = 1, 2, \dots, |R| \text{ tq } k \neq i}_{\Theta(n)} \left\{ \begin{array}{ll} V_i[k] + 1 \wedge F_i[k] \leftarrow score(A_i, A_k), & \text{si } A_k \text{ es vecino de } A_i \\ V_i[k] \leftarrow 0 \wedge F_i[k] \leftarrow -1, & \text{de lo contrario} \end{array} \right. \underbrace{\hspace{10em}}_{\Theta(k) + \Theta(score(A_i, A_k))}$$

$$\underbrace{A_p \leftarrow \max(F_i[1], \dots, F_i[|R|])}_{\Theta(n)}$$

$\rightarrow A_p$

Por lo tanto, la complejidad temporal de esta sección del algoritmo se define de la siguiente manera:

$$\begin{aligned}
 \therefore \Theta(\text{peerfs}(A_i)) &= \Theta(n) * (\Theta(k) + \Theta(\text{score}(A_i, A_k))) + \Theta(n) \\
 &\quad \Theta(n * k) + \Theta(n * \text{score}(A_i, A_k)) + \Theta(n) \\
 &\quad \Theta(n) + \Theta(n * \text{score}(A_i, A_k) + \Theta(n)) \tag{3.2} \\
 &\quad \Theta(2n) + \Theta(n * \text{score}(A_i, A_k)) \\
 &\quad \Theta(n) + \Theta(n * \text{score}(A_i, A_k))
 \end{aligned}$$

Por consiguiente, queda pendiente analizar la complejidad del algoritmo de puntaje:

$$\underbrace{\text{score}(A_i, A_k) = 0.5 \cdot e^{-4 \cdot e^{-0.5V_i[k]}} + 0.5 \cdot e^{-0.1 \cdot e^{0.5E(A_k, A_i)}}}_{\Theta(k) + \Theta(k)}$$

$$\begin{aligned}
 \therefore \Theta(\text{score}(A_i, A_k)) &= \Theta(k) + \Theta(k) \\
 &= \Theta(2 * k) \tag{3.3} \\
 &= \Theta(k)
 \end{aligned}$$

Al completar (3.2) con (3.3) se obtiene que:

$$\begin{aligned}
 \therefore \Theta(\text{peerfs}(A_i)) &= \Theta(n) * (\Theta(k) + \Theta(\text{score}(A_i, A_k))) + \Theta(n) \\
 &\quad \Theta(n * k) + \Theta(n * \text{score}(A_i, A_k)) + \Theta(n) \\
 &\quad \Theta(n) + \Theta(n * \text{score}(A_i, A_k) + \Theta(n)) \\
 &\quad \Theta(2n) + \Theta(n * \text{score}(A_i, A_k)) \\
 &\quad \Theta(n) + \Theta(n * \text{score}(A_i, A_k)) \tag{3.4} \\
 &\quad \Theta(n) + \Theta(n * k) \\
 &\quad \Theta(n) + \Theta(n) \\
 &\quad \Theta(2n) \\
 &\quad \Theta(n)
 \end{aligned}$$

Y finalmente se completa (3.1) con (3.4):

$$\begin{aligned}
 \therefore \Theta(\text{peer}(A_i)) &= \Theta(k) + \Theta(\text{peerfs}(A_i)) \\
 &= \Theta(k) + \Theta(\text{peerfs}(A_i)) \\
 &= \Theta(k) + \Theta(n) \\
 &= \Theta(k * n) \\
 &= \Theta(n)
 \end{aligned} \tag{3.5}$$

En una red no distribuida, todos los cálculos se ejecutarían en un único punto, con lo cual se obtiene:

$$\underbrace{\underbrace{\forall A_i \in R}_{\Theta(n)} \underbrace{\text{peer}(A_i)}_{\Theta(n)}}_{\Theta(n^2)}$$

Sin embargo, al tratarse de un ambiente distribuido, el cálculo de par lo hace cada agente [48], para dejar el tiempo final en:

$$\underbrace{\underbrace{1}_{\Theta(1)} \underbrace{\text{peer}(A_i)}_{\Theta(n)}}_{\Theta(n * 1) = \Theta(n)}$$

Por lo tanto, la complejidad temporal para el algoritmo es lineal.

Diseño de los experimentos

“This progress among the people through Dhamma has been done by two means, by Dhamma regulations and by persuasion. Of these, Dhamma regulation is of little effect, while persuasion has much more effect” .

King Ashoka

En este punto de la investigación se necesita corroborar la veracidad de la hipótesis plateada en el capítulo 1; para ello se ha elegido el modelo de diseño estadístico de experimentos explicado en [37] y [58]. La principal razón para esta escogencia es la gran aceptación y uso extendido en gran cantidad de proyectos científicos.

El diseño de experimentos (DoE, por sus siglas en inglés) es un enfoque sistemático y riguroso para la resolución de problemas ingenieriles que aplica principios y técnicas a una colección de datos, con el fin de asegurar una conclusión válida, defendible y soportable. Además de lo anterior, este modelo es llevado a cabo bajo la condición de minimizar el gasto de recurso, corridas de los experimentos, tiempo y dinero [38].

Existen cuatro tipos de problemas generales en los cuales es provechoso aplicar diseño de experimentos:

1. **Comparativa:** el ingeniero está interesado en saber si el cambio de un factor genera mejora en un proceso.
2. **Caracterización:** se desea saber cuáles son los factores que afectan un proceso, para poder ordenarlos según su nivel de impacto.
3. **Modelado:** es de interés modelar un proceso cuya salida sea una función matemática con alto poder predictivo y tener un buen estimado de los coeficientes de dicha función.
4. **Optimización:** el ingeniero desea determinar los parámetros óptimos de los factores de un proceso. Esto es, determinar para cada factor el nivel que optimiza la respuesta del proceso.

La presente investigación diseñará una serie de experimentos que permitan optimizar el comportamiento de la selección de pares basada en criterio previsor y comparar el efecto de esta sobre el uso de los otros criterios de redes convencionales. Por lo tanto, la metodología de diseño de experimentos permitió generar una caracterización para la cantidad de desconexiones.

4.1 Factores y niveles

En el diseño de experimentos, un *factor* es aquel componente que tiene cierta influencia en las variables de respuesta [37]. El objetivo de un experimento es determinar esta influencia. A su vez, cada factor cuenta con varios *niveles* posibles con los cuales experimentar.

Con base en la información recolectada en esta etapa de la investigación, así como en la experiencia adquirida por el autor y expuesta en los capítulos anteriores, se han seleccionado los siguientes factores y niveles para su estudio:

1. Criterio de selección
 - (a) Tiempo
 - (b) Señal
 - (c) Aleatorio
 - (d) Previsor
2. Tiempo de vida de la red
 - (a) 1000 pasos
 - (b) 500 pasos
 - (c) 10 000 pasos
 - (d) 15 000 pasos
3. Población
 - (a) Pequeña
 - (b) Mediana
 - (c) Grande
4. Patrón de movimiento

4.1. Factores y niveles

- (a) Convergencia/divergencia
- (b) Concentración
- (c) Escuadrón en movimiento
- (d) Aleatorio

Potencialmente, la cantidad de factores que influyen en las variables de respuesta es infinita; sin embargo, por razones de tiempo y presupuesto se han seleccionado las que se considera ejercen mayor influencia en las variables de respuesta (Alcances y limitaciones), tal y como se resume en el cuadro 4.1. Algunos factores que no se considerarán en esta investigación, pero que también pueden ejercer influencia, son:

- El patrón de radiación de la antena
- El fabricante o modelo del dispositivo
- La topología del terreno
- Plataforma de implementación (*hardware vs. software*)
- La velocidad de desplazamiento
- El medio ambiente

		Factores			
		Criterio	Tiempo	Población	Migración
Niveles	Señal	1000	Pequeña	Convergencia	
	Tiempo	5000	Mediana	Concentración	
	Azar	10 000	Grande	Escuadrón	
	Previsor	15 000		Aleatorio	

Table 4.1: Los factores y sus respectivos niveles analizados en esta investigación.

4.2 Variables de respuesta

Dado que la hipótesis propone minimizar las desconexiones, solamente es necesario un contador para dichos fallos (ver Tolerancia a fallos). La selección de este valor como variable de respuesta tiene varios puntos a su favor:

1. Un número entero positivo. No es necesaria ninguna estructura de datos compleja, por lo que la simulación puede darse en cualquier lenguaje de programación moderno.
2. Llevar los datos de la simulación al análisis estadístico se vuelve muy sencillo, pues solamente se requiere guardar los datos en un archivo de texto plano.
3. Similar a la simulación, el paquete estadístico puede manejar enteros positivos sin ningún problema.
4. Fácil de interpretar al ojo humano, pues será mejor mientras haya menos desconexiones.

4.3 Recolección de datos

Como ya se mencionó, el simulador genera los datos para cada paso de la ejecución. Mediante *scripts*, estos datos fueron ordenados de tal forma que puedan ser usados como entrada para el paquete estadístico R [43], el cual permite realizar un análisis de experimentos factoriales [58].

4.4 Análisis de varianza

El análisis de varianza (ANOVA, por sus siglas en inglés) permite asegurar que la variación en los resultados de un experimento no es mayor a la suma de variaciones de los factores y un cierto grado de error en sus medidas. Esto permite aceptar o

rechazar la hipótesis con una probabilidad de error (de preferencia muy baja), con base en evidencia estadística [58].

El ANOVA tiene como objetivo analizar la relación entre una variable cuantitativa X y una variable cualitativa Y de k atributos. Cada atributo i define una población dada por la variable cuantitativa [46].

X_i : variable X restringida al atributo i .

Así, se tienen k poblaciones X_1, X_2, \dots, X_n (llamadas tratamientos) que se suponen normales, independientes, con varianzas similares y con medias poblacionales $\mu_1, \mu_2, \dots, \mu_n$. Se desea determinar si X no varía según el atributo de Y , es decir, si las poblaciones son equivalentes y entonces los tratamientos son igualmente efectivos. Para ello se plantean y contrastan las hipótesis:

H_0 : X no varía según el atributo de Y
(poblaciones equivalentes, es decir, $\mu_1 = \mu_2 = \dots = \mu_k$).

H_1 : X varía según el atributo de Y
(poblaciones no equivalentes), al menos dos de las medias no son iguales.

Un análisis de varianza permite saber si la diferencia en la media de varias poblaciones es significativa debido a la influencia de alguno de los factores. Una gran cualidad de ANOVA es que permite analizar varias poblaciones, mientras que otros métodos no permiten más de dos.

Aunque existen otras opciones como SPSS® y Minitab®, para el análisis de los resultados de esta investigación se decidió utilizar R. Entre las razones por las cuales se tomó dicha determinación están:

- Presupuesto: R es de acceso gratuito.
- Filosofía: al ser *software* libre, va de acuerdo con la filosofía del autor principal de esta obra.
- Uso en línea de comandos: facilitó la creación de los *scripts* para automatizar los experimentos, recolección de datos y, finalmente, el análisis de resultados.
- Facilidad de herramientas: no solo tiene una interfaz gráfica, sino que hay varias interfaces amigables que facilitaron el análisis de casos específicos.

Resultados

“A los demonios no hay que creerles
ni cuando dicen la verdad”.

Gabriel García Márquez

5.1 Experimentos

Con base en la metodología de DoE [37] se realizaron 3840 ejecuciones del experimento. Este número es el resultado de la multiplicación de los factores descritos en la tabla 4.1.

$$\begin{aligned} \text{Ejecuciones} &= \text{Criterio} * \text{Tiempo} * \text{Población} * \text{Migración} * \text{Réplicas} \\ &= 4 * 4 * 3 * 4 * 20 \\ &= 3840 \end{aligned}$$

Una muestra de los resultados sin analizar (solo lo producido por el simulador) se encuentra en el anexo E.

5.2 Supuestos de ANOVA

Antes de proceder a evaluar los resultados con el modelo de ANOVA, estos deben cumplir ciertos criterios para asegurar que las respuestas son confiables. De acuerdo con [4] [37], las diferencias en las varianzas de los residuos deben ser:

- **Independientes:** la ocurrencia de un experimento no debe afectar los resultados de ningún otro.
- **Normales:** siguen una distribución normal.
- **Homogéneas:** esta propiedad también es conocida como homocedasticidad. Dicha condición se cumple si al formar los resultados en grupos, estos tienen la misma varianza.

Para más información acerca de estos supuestos el lector puede consultar [35].

El supuesto de independencia se asegura mediante una ejecución aleatoria de los experimentos [4] [37].

R proporciona herramientas visuales para comprobar los demás supuestos:

- **Gráfico de ubicación-escala:** usado para medir que la distribución sea homogénea. No deben apreciarse patrones ni agrupamientos.
- **Histograma de residuos:** usado para analizar una distribución normal. Las barras deben estar debajo de la curva de una distribución normal lo más que se pueda.
- **Gráficos Q-Q:** también usados para verificar una distribución normal. Los puntos deben estar apegados a la línea.

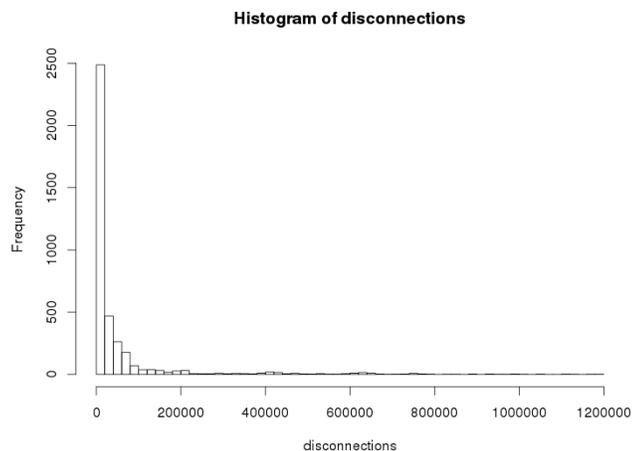


Figura 5.1: Histograma para la frecuencia de las desconexiones.

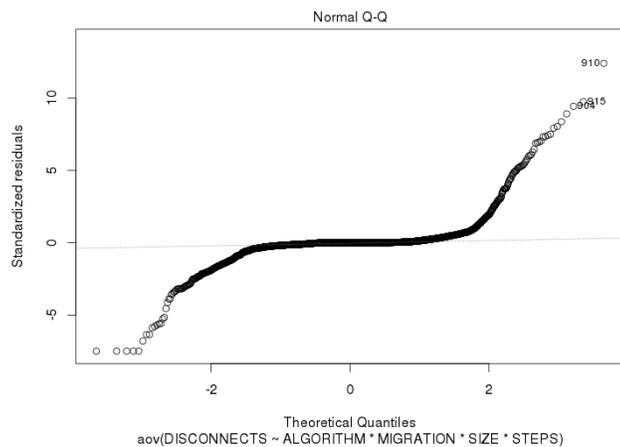


Figura 5.2: Gráfico Q-Q para el análisis de normalidad en la varianza de las desconexiones.

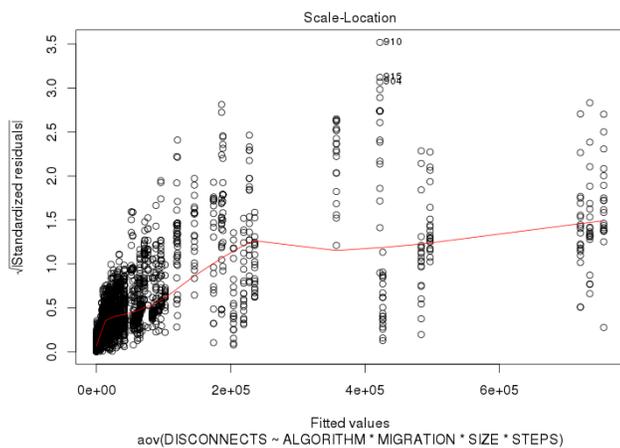


Figura 5.3: Gráfico de ubicación escala para la varianza de las desconexiones.

En la figura 5.1 se puede apreciar que las varianzas se alejan por mucho de una distribución normal, y se corrobora con el alejamiento de los puntos en el gráfico Q-Q (figura 5.2). Si se agrega que los residuos no son homogéneos (figura 5.3), se puede concluir que los datos no son aptos para un análisis de varianza.

Cuando la variable de respuesta es una tasa no proporcional a los factores, es muy común encontrar un gráfico como el de la figura 5.1 [4].

Cuanto mayor sea la densidad de la red, y cuanto más tiempo esta se encuentre bajo evaluación, es de esperar que crezca la cantidad de errores. Ajustar la variable de respuesta para que sea proporcional a dos factores es un paso complicado y podría generar muchos vicios en el análisis de los resultados. Para efectos de rechazar o aceptar la hipótesis propuesta en 1.1, el análisis debería centrarse en los efectos de los algoritmos y no en los del tiempo o tamaño de la red (aunque sí son de interés).

5.2.1 Transformaciones Box-Cox

Ante situaciones como esta, es muy común aplicar otro tipo de transformaciones a la variable de respuesta [37]. Los estadísticos George Box y David Cox desarrollaron un procedimiento para identificar un exponente apropiado ($\Lambda = 1$) que se utilice para transformar los datos en una ‘forma normal’. El valor λ indica la potencia a la que deben plantearse todos los datos. Para ello, la transformación de potencia de Box-Cox busca desde $\Lambda = -5$ a $\Lambda = 5$, hasta que se encuentre el mejor valor [11].

Lambda	Variable de respuesta
-2	$Y^{-2} = 1/Y^2$
-1	$Y^{-1} = 1/Y^1$
-0.5	$Y^{-0.5} = 1/\sqrt{Y}$
0	$\log(Y)$
0.5	$Y^{0.5} = 1/\sqrt{Y}$
1	Y^1
2	Y^2

Table 5.1: Tabla de transformaciones Box-Cox.

De acuerdo con las transformaciones propuestas en [10] y resumidas en el cuadro 5.1, se aplicaron las transformaciones y se determinó (según el criterio del autor) que la transformación logarítmica (Lambda) cumplía con los requisitos de ANOVA.

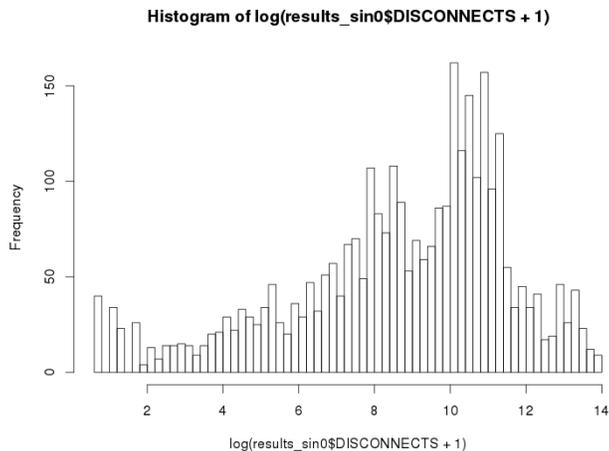


Figura 5.4: Histograma para la frecuencia del logaritmo de las desconexiones.

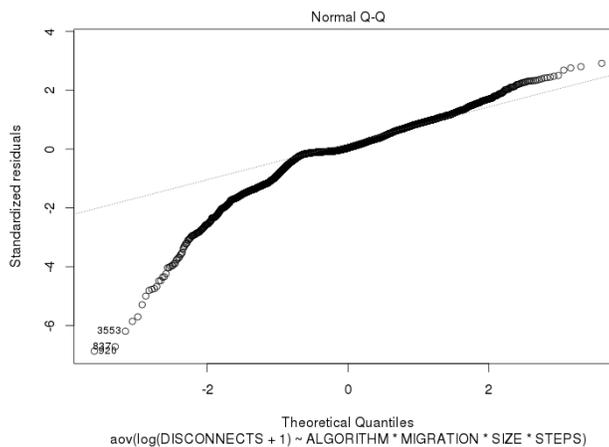


Figura 5.5: Gráfico Q-Q para el análisis de normalidad en la varianza del logaritmo de las desconexiones.

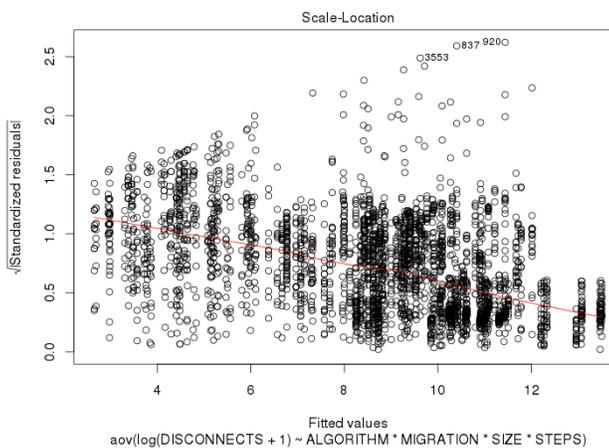


Figura 5.6: Gráfico de ubicación escala para la varianza del logaritmo las desconexiones.

Aunque los resultados del histograma (figura 5.4) se asemejan al área bajo la curva de una distribución normal, se puede apreciar un pico en la parte derecha que se aleja bastante, así como una depresión en lo que debería ser la parte alta de la curva.

En el gráfico ubicación escala (figura 5.5) se presenta una concentración en la parte inferior derecha, mientras que la representación Q-Q (figura 5.5) manifiesta una cola pronunciada en la parte inferior izquierda.

Aunque la visualización de los datos ya da indicios de que la distribución de los residuos no sigue una distribución normal, no fue sino hasta que se aplicó un test de Tukey (función `tukeyHSD` en R) cuando se comprobó que los datos no eran aptos.

5.2.2 ANOVA sobre una distribución Weibull

La cola que se encuentra en el gráfico Q-Q (figura 5.5) es muy interesante, pues es común en conjuntos de mediciones para el fallo de máquinas [41]. En estos casos, los resultados se pueden analizar al utilizar una distribución de Weibull.

Los paquetes *fitdistrplus* y *survival* para R permiten la manipulación de resultados bajo una distribución Weibull, para su procesamiento con las funciones de ANOVA. En la figura 5.7 se muestra la aplicación de las herramientas visuales, mientras que en la figura 5.8 se utiliza un test formal para constatar la aptitud de los datos.

Así las cosas, el logaritmo de la cantidad de desconexiones analizado sobre una distribución Weibull apto para analizar su varianza se muestra de la siguiente manera.

```
> fitdist(v_response, "weibull")
Fitting of the distribution 'weibull' by maximum likelihood
Parameters:
      estimate Std. Error
shape      4.28    0.0633
scale      9.91    0.0439
```

Figura 5.8: Resultados de la función *fitdist* para verificar que los datos pueden ser procesados por ANOVA. Con que los resultados sean mayores o cercanos a 0.5 se puede concluir que siguen una distribución Weibull.

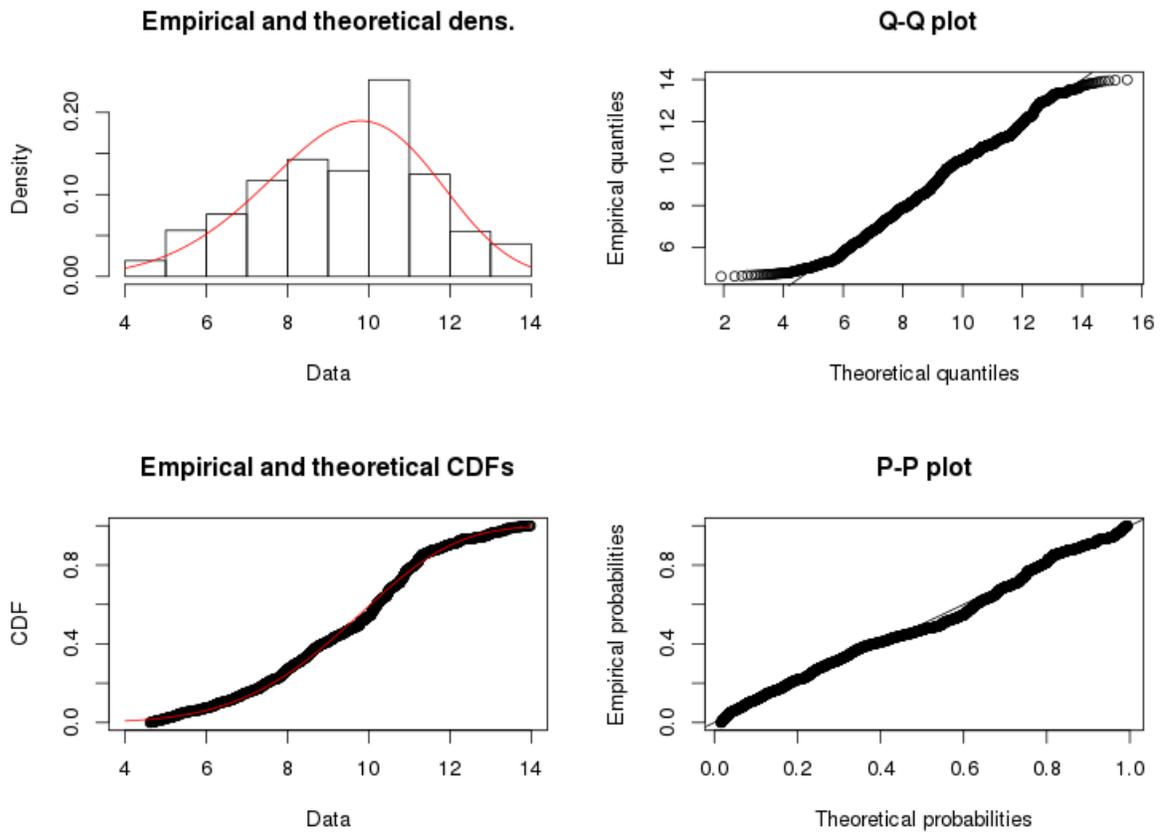


Figura 5.7: Herramientas gráficas para la valoración visual de la normalidad en la variable de respuesta bajo una distribución Weibull. En la imagen se aplican a la varianza del logaritmo de las desconexiones.

5.3 Análisis de varianza

Siempre de acuerdo con la metodología de ANOVA, tal como se explica en [4] [37], el paso por seguir es analizar cuál es la probabilidad de que el logaritmo natural de la cantidad de desconexiones sea afectado por algún factor o la combinación de alguno de ellos. La función *summary* de R muestra los resultados.

FACTOR	Df	Deviance	Resid. Df	-2*LL	Pr(>Chi)	Sig
ALGORITHM	3	571.79	2839	11660	1.31e-123	99.9%
MIGRATION	3	9.18	2836	11651	2.70e-02	<95%
SIZE	2	1861.56	2834	9790	0.00e+00	99.9%
STEPS	3	2175.68	2831	7614	0.00e+00	99.9%
ALGORITHM:MIGRATION	9	4.46	2822	7609	8.79e-01	<95%
ALGORITHM:SIZE	6	27.35	2816	7582	1.24e-04	99.9%
MIGRATION:SIZE	6	25.61	2810	7557	2.63e-04	99.9%
ALGORITHM:STEPS	9	24.81	2801	7532	3.19e-03	99.9%
MIGRATION:STEPS	9	7.30	2792	7524	6.06e-01	<95%
SIZE:STEPS	6	25.50	2786	7499	2.76e-04	95%
ALGORITHM:MIGRATION:SIZE	18	15.87	2768	7483	6.02e-01	<95%
ALGORITHM:MIGRATION:STEPS	27	21.12	2741	7462	7.81e-01	<95%
ALGORITHM:SIZE:STEPS	18	23.58	2723	7438	1.69e-01	99%
MIGRATION:SIZE:STEPS	18	17.47	2705	7421	4.91e-01	95%
ALGORITHM:MIGRATION:SIZE:STEPS	54	34.65	2651	7386	9.81e-01	<95%

Table 5.2: Resultados de R para el análisis de varianza en el logaritmo natural de las desconexiones. $Pr(> Chi)$ es la probabilidad de que las varianzas sean iguales, mientras que la columna Sig representa la certeza estadística de que la combinación de factores tenga un efecto sobre la variable de respuesta.

La figura 5.2 muestra las combinaciones de factores que generan un impacto. La columna $Pr(> Chi)$ establece la probabilidad de que H_0 en la ecuación 4.4 sea cierta, es decir, que las varianzas sean iguales. En caso de ser menor a 0.5, el resultado se puede interpretar como que H_0 es falsa y H_1 es verdadera, lo cual indica que la combinación sí tiene un efecto sobre el logaritmo de las desconexiones (la probabilidad de H_1 es igual a $1 - P[H_0]$).

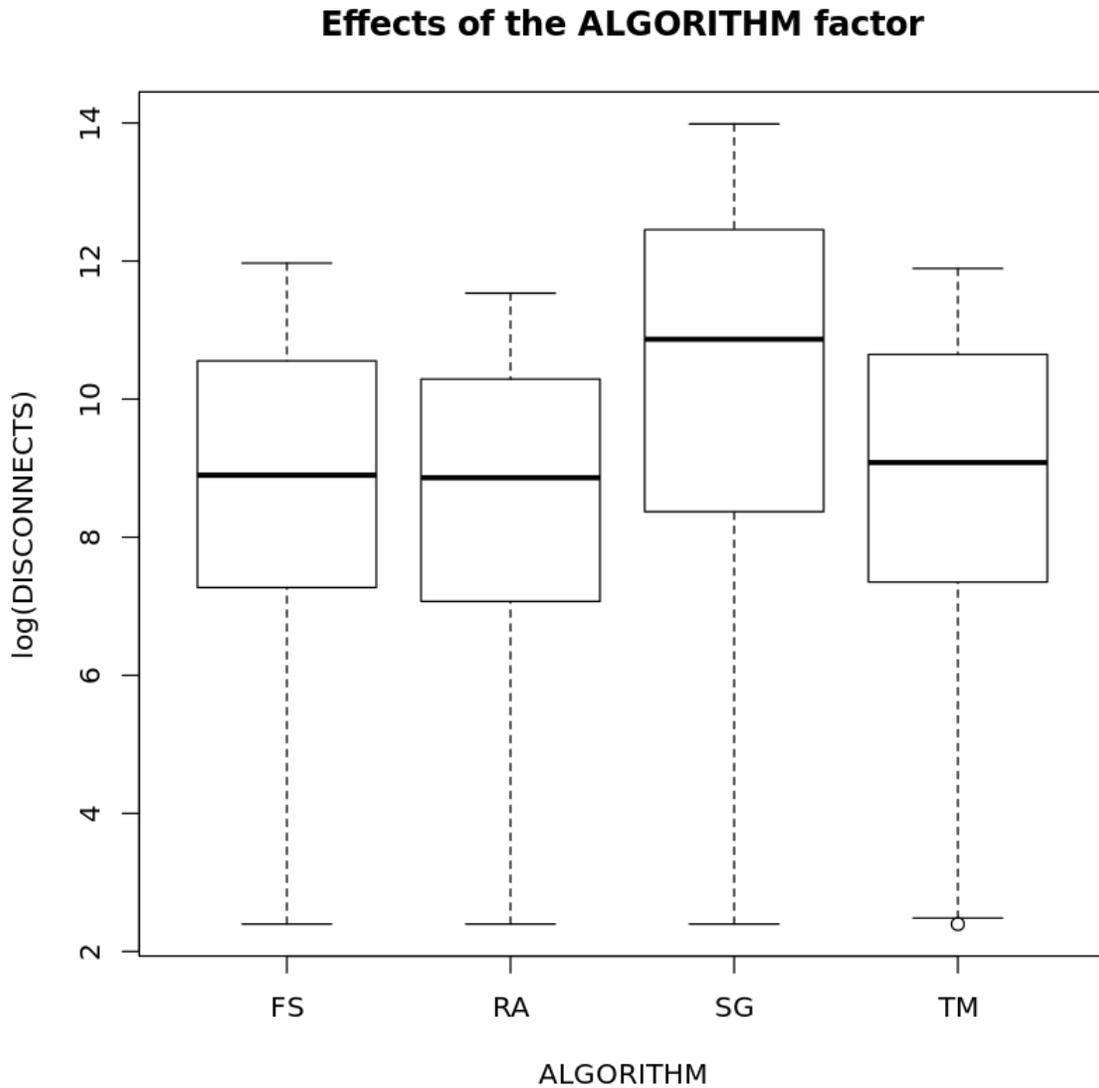


Figura 5.9: Representación gráfica de los efectos del factor algoritmo sobre el logaritmo natural de las desconexiones.

Así pues, la primera línea muestra que el algoritmo sí tiene un impacto sobre la variable de respuesta, por lo que es necesario investigar cuál es su magnitud y consecuencias sobre el modelo. La figura 5.9 corresponde al análisis de esta combinación específica (de un solo factor) donde se establece su efecto. Sin embargo, la diferencia entre los algoritmos predictor y aleatorio no es fácilmente distinguible, por lo que requerirá un estudio más detallado.

En el cuadro 5.3 el lector podrá encontrar las combinaciones que generan un efecto con una certeza estadística mayor al 95 %.

Combinación	Significancia	Referencias
Algoritmo	99.9%	figura 5.9 y cuadro 5.4
Tamaño	99.9%	figura 5.10 y cuadro 5.5
Pasos	99.9%	figura 5.11 y cuadro 5.6
Algoritmo y tamaño	99.9%	figura 5.12
Migración y tamaño	99.9%	figura 5.13
Algoritmo y pasos	99.9%	figura 5.15
Algoritmo, tamaño y pasos	99%	figuras 5.17,5.18 y 5.19
Tamaño y pasos	95%	figura 5.16
Migración, tamaño y pasos	95%	figuras 5.20,5.21 y 5.22

Table 5.3: Combinaciones de factores que generan un efecto no aleatorio sobre el modelo.

5.3.1 Factor algoritmo

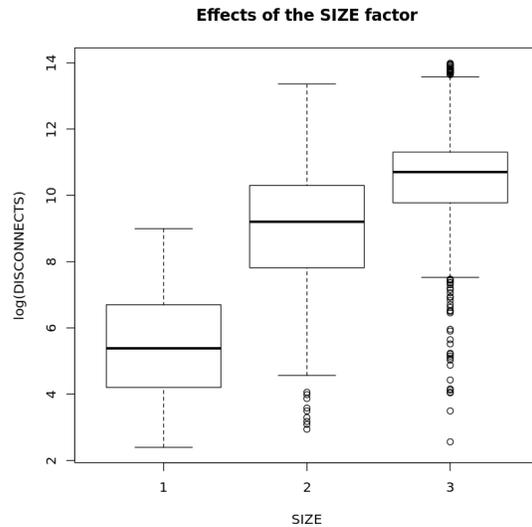
ALGORITHM				
	FS	RA	SG	TM
	-0.314	-0.6	1.12	-0.226
rep	796	793	808	800

Table 5.4: Análisis de efectos para el factor algoritmo. Significancia sobre el 99.9 %.

En el cuadro 5.4 se muestran los datos concretos para el efecto del algoritmo sobre la variable de respuesta. Los números representan el estado de la media con respecto a la media global; así, un número negativo indica que el factor tiene una media por debajo de la global (menor cantidad de desconexiones). En el caso del factor algoritmo,

el criterio predictor tiene un valor menor a los de señal y tiempo, pero no menor al aleatorio. El estudio de las otras combinaciones revelará una mejor interpretación de los resultados, lo cual se verá más adelante en este documento.

5.3.2 Factor tamaño



		SIZE		
		1	2	3
		-4.06	0.43	1.87
rep		685	1254	1258

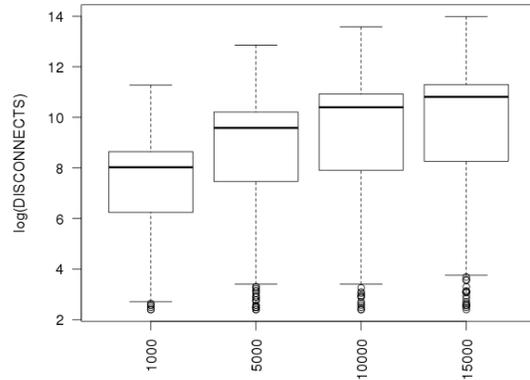
Table 5.5: Análisis de efectos para el factor tamaño. Significancia sobre el 99.9 %.

Figura 5.10: Representación gráfica de los efectos del factor tamaño sobre el logaritmo natural de las desconexiones.

En el cuadro 5.5 puede apreciarse que el logaritmo de la cantidad de desconexiones crece conforme se agregan más agentes a la simulación. El resultado es esperado, pues la cantidad de desconexiones no es una tasa proporcional al factor, tal como se había explicado al inicio de este capítulo.

No obstante, es de interés la cantidad de valores atípicos (también conocidos como *outliers*) que se encuentran por debajo de la media. Nuevamente, se requiere más estudio para explicar este comportamiento.

5.3.3 Factor pasos



	SIZE		
	1	2	3
	-4.06	0.43	1.87
rep	685	1254	1258

Table 5.6: Análisis de efectos para el factor pasos. Significancia sobre el 99.9 %.

Figura 5.11: Representación gráfica de los efectos del factor tamaño sobre el logaritmo natural de las desconexiones.

Al igual que el factor tamaño, en el cuadro 5.6 y en la figura 5.11 el logaritmo de la cantidad de desconexiones crece mientras la simulación sea más extendida. Al no ser proporcionado a este factor, el resultado es esperado.

5.3.4 Factores algoritmo y tamaño

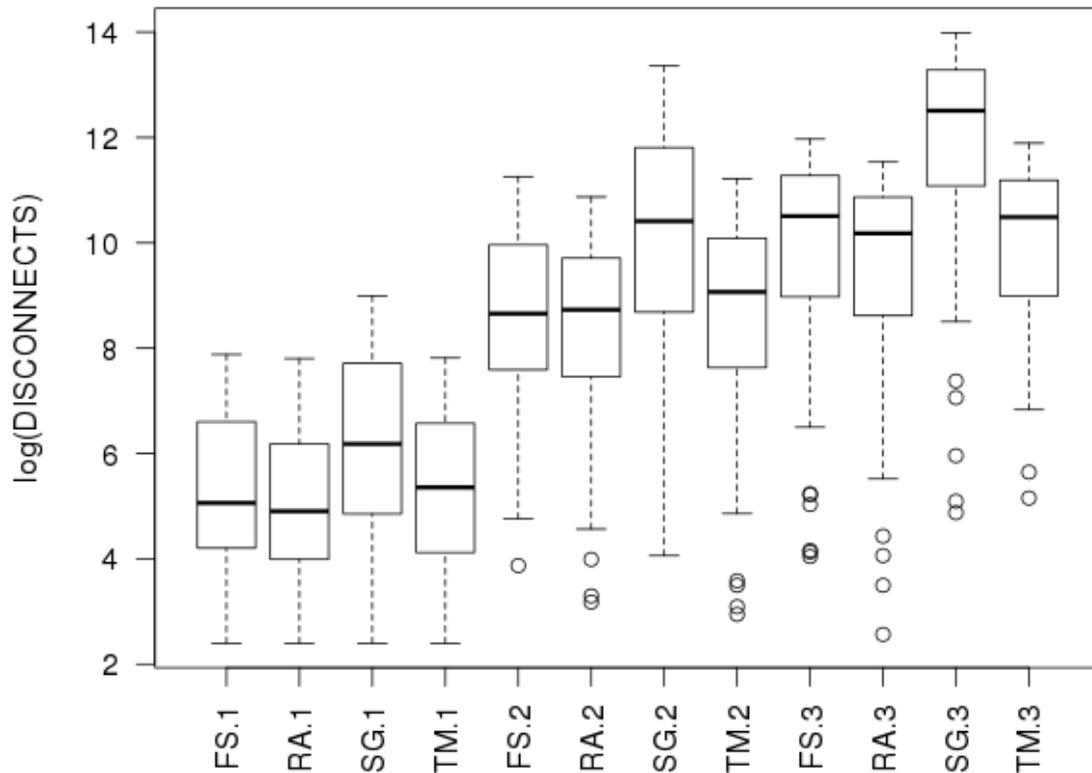


Figura 5.12: Análisis de efectos para los factores algoritmo y tamaño. Significancia sobre el 99.9%.

Al combinar los factores de algoritmo y tamaño (figura 5.12) se obtiene que el algoritmo aleatorio sigue teniendo un logaritmo de la cantidad de desconexiones menor al predictor.

Se encuentra que el criterio predictor es superior al aleatorio en poblaciones de agentes medianamente densas. La ventaja que el algoritmo aleatorio le saca al predictor en redes de tamaño 1 y 3 hace que parezca ser el ganador absoluto al estudiar la figura 5.9 (factor algoritmo).

5.3.5 Factores migración y tamaño

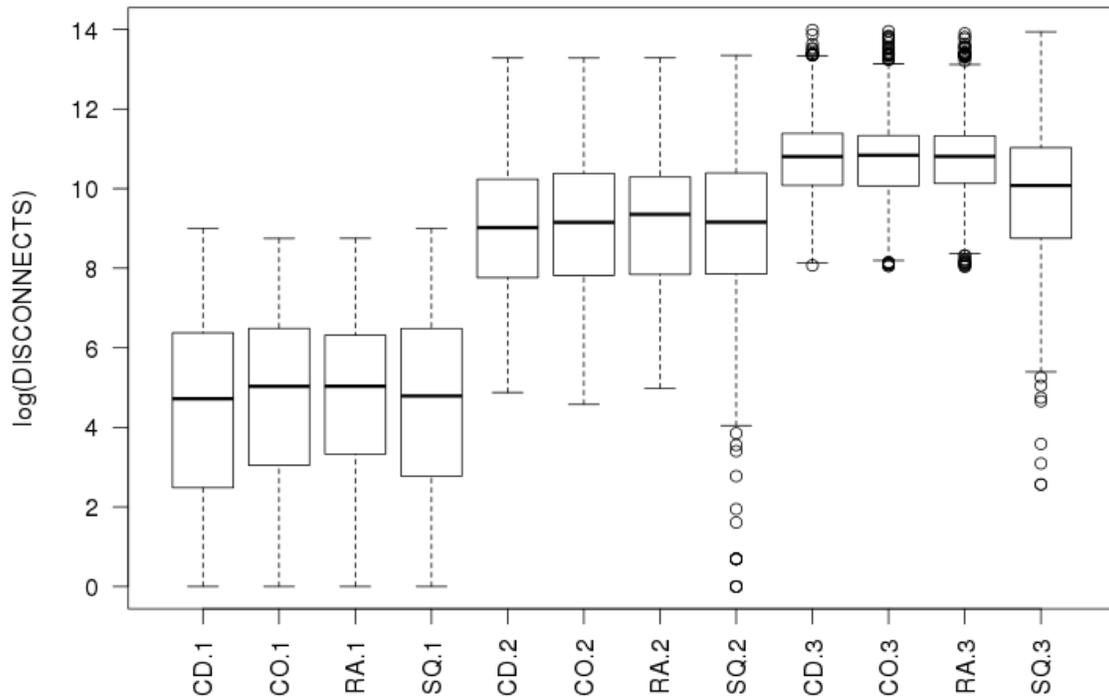


Figura 5.13: Análisis de efectos para los factores migración y tamaño. Significancia sobre el 99.9 %.

El estudio de esta interacción de factores muestra que el modelo de escuadrón tiene un mejor resultado (más cercano a cero) en todos los tamaños utilizados durante la simulación. Además, presenta más muestras atípicas por debajo de la media cuando la red de agentes se hace más densa. En la figura 5.13 se presentan los resultados antes descritos, que también encajan con las expectativas del autor, pues bajo un patrón de migración se espera que los agentes se mantengan juntos por más tiempo durante su recorrido.

Es interesante mencionar que, con excepción de las combinaciones con el patrón de escuadrón, las redes de tamaño 3 generan muchos valores atípicos. Con base en la

observación de estas simulaciones se encontró que los escenarios estaban hiperconectados, donde los agentes casi siempre podían encontrar al menos un par dentro de la red.

Lo anterior hace que muchas simulaciones tuvieran buenos resultados. Sin embargo, y también debido a la hiperconectividad, también se encontraron muchas simulaciones en las que estos agentes encontraron un par en algún momento, pero lo perdían rápidamente.

Este último caso se repetía durante el tiempo de vida y generaba los valores atípicos encontrados (figura 5.14). Si la red fuera menos densa, esos agentes habrían quedado fuera de la red (huérfanos) durante toda la simulación.

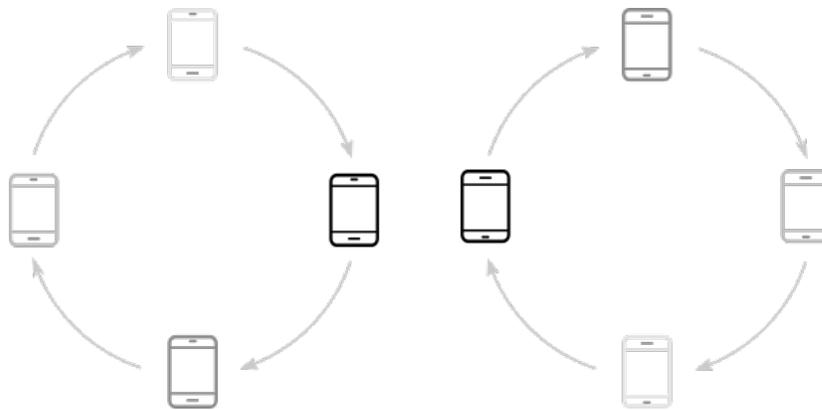


Figura 5.14: Comportamientos interesantes al tratar poblaciones de agentes muy densas.

5.3.6 Factores algoritmo y pasos

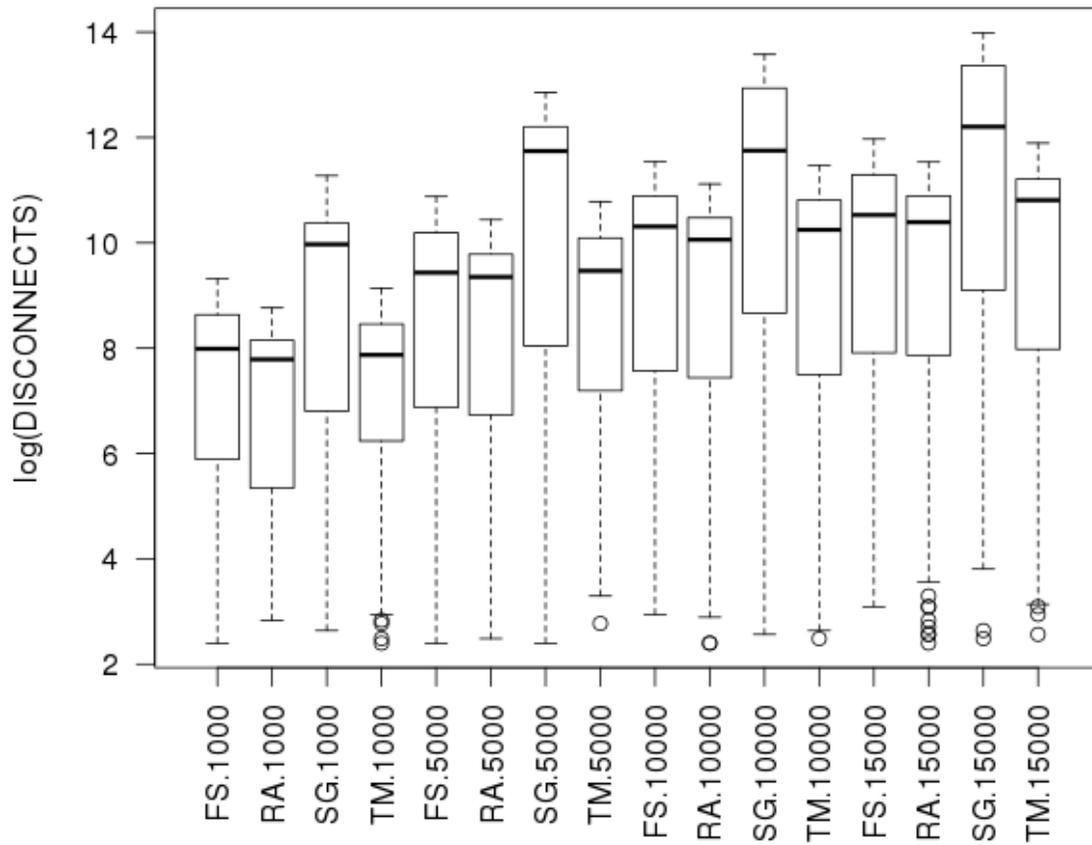


Figura 5.15: Análisis de efectos para los factores algoritmo y pasos. Significancia sobre el 99.9 %.

5.3.7 Factores tamaño y pasos

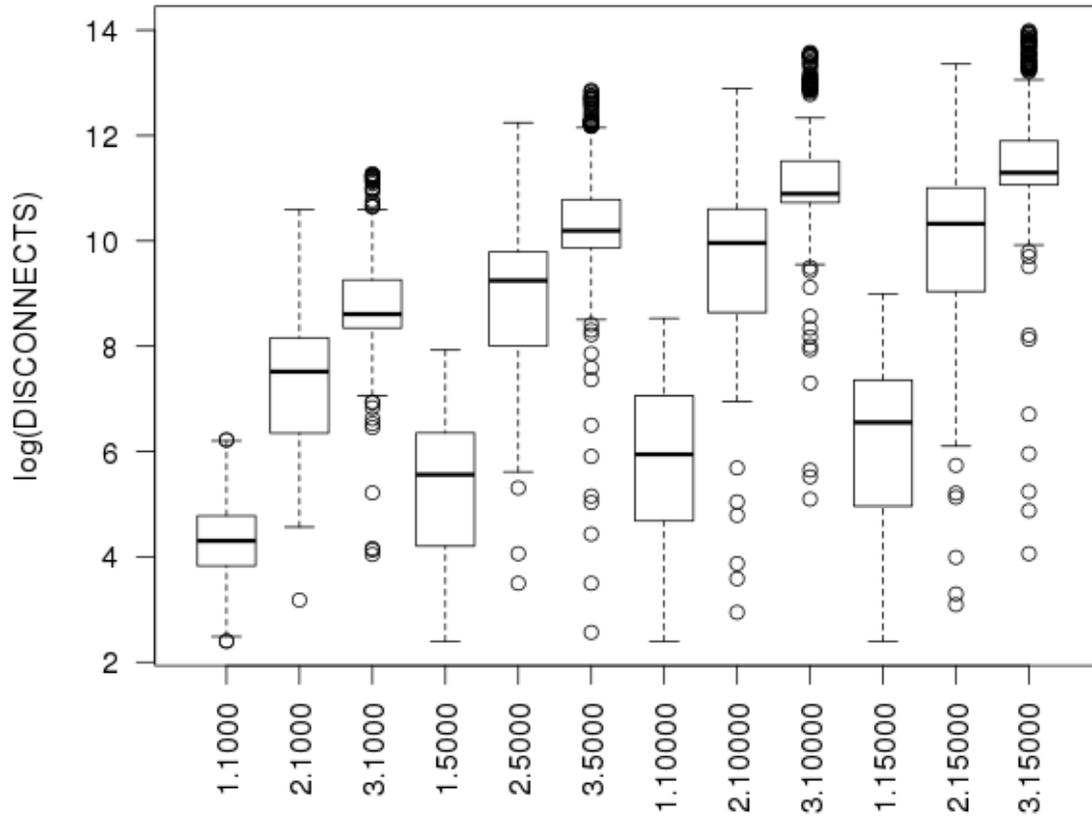


Figura 5.16: Análisis de efectos para los factores tamaño de la red, patrón de migración, algoritmo y tiempo de la vida de la red. Significancia sobre el 99.9 %.

5.3.8 Factores algoritmo, tamaño y pasos

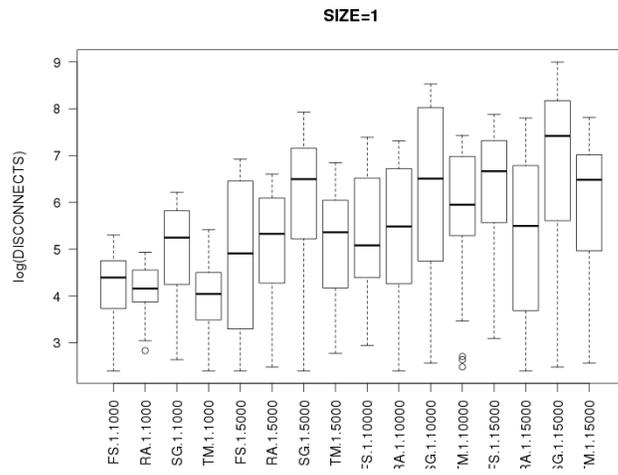


Figura 5.17: Interacción de los factores algoritmo y tiempo para poblaciones de tamaño nivel 1.

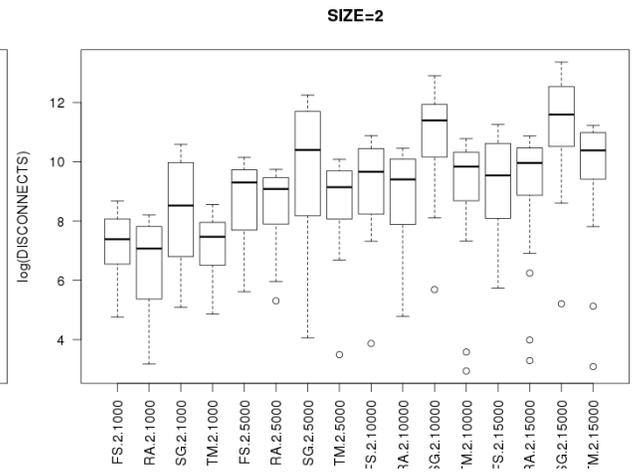


Figura 5.18: Interacción de los factores algoritmo y tiempo para poblaciones de tamaño nivel 2.

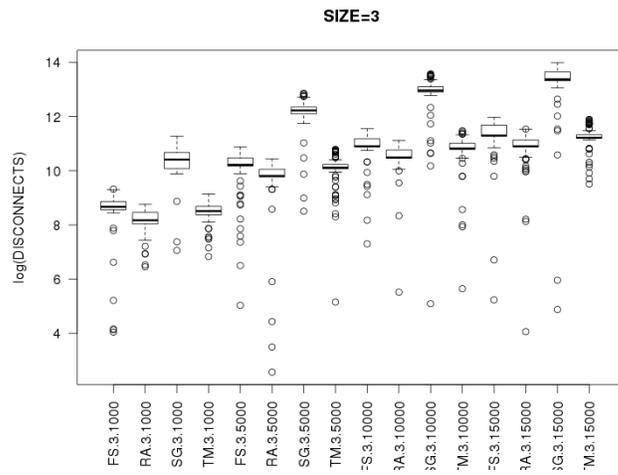


Figura 5.19: Interacción de los factores algoritmo y tiempo para poblaciones de tamaño nivel 3.

En las figuras 5.17 y 5.18 se encuentra que el algoritmo predictor tiene mejores resultados que los de azar, tiempo o señal. Sin embargo, en la figura 5.19 el algoritmo aleatorio tiene un mejor valor (cercano a cero) en todas las combinaciones.

5.3.9 Factores migración, tamaño y pasos

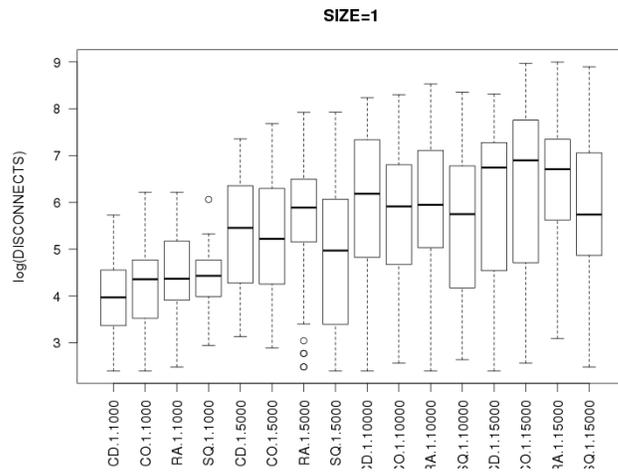


Figura 5.20: Interacción de los factores migración y tiempo para poblaciones de tamaño nivel 1.

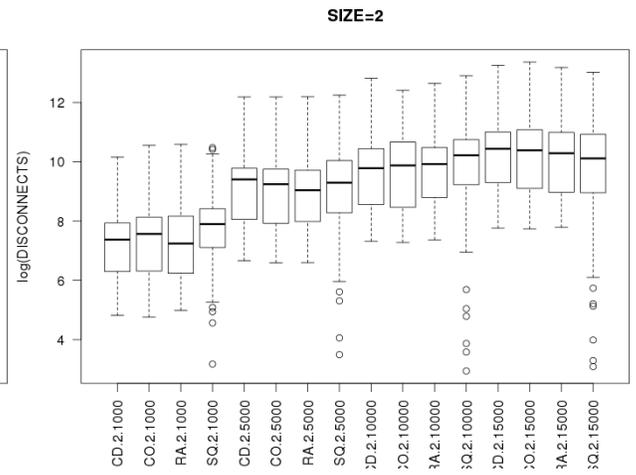


Figura 5.21: Interacción de los factores migración y tiempo para poblaciones de tamaño nivel 2.

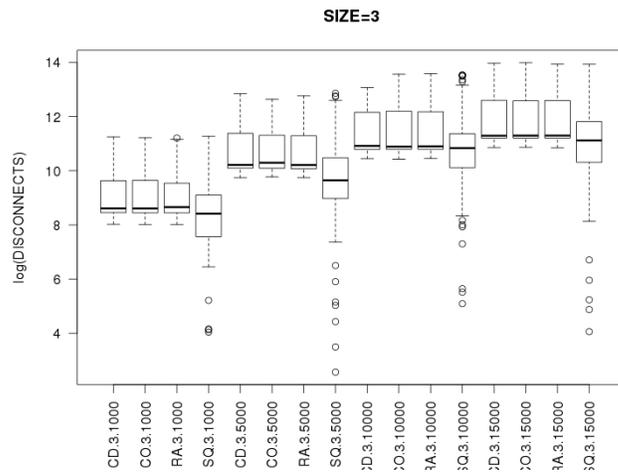


Figura 5.22: Interacción de los factores migración y tiempo para poblaciones de tamaño nivel 3.

Resulta interesante que el patrón de escuadrón no tiene tan buenos resultados (cerca de cero) en escenarios medianamente densos (5.21). En las figuras 5.20 y 5.22 se aprecia un comportamiento más parecido a las otras combinaciones estudiadas.

5.3.10 Resumen

A modo de resumen, el autor de este trabajo propone los siguientes puntos como resultados de interés:

- Los datos no siguen una distribución normal.
- Después de una transformación, se generó un set de datos cuyo diagrama Q-Q muestra una cola a su lado derecho.
- Este comportamiento es característico de los datos que representan errores en una máquina.
- El análisis de varianza se realizó bajo una distribución Weibull.
- El criterio aleatorio presenta mejores resultados por su ventaja en escenarios densos.
- Los escenarios densos presentan muchas muestras atípicas. La observación de las simulaciones concluyó que la hiperconexión es la causa.

Conclusiones y trabajo futuro

“Luckily, pure science had a friend and patron in its wealthier and better-looking cousin, technology. Technology was the branch of human knowledge by which the theoretical became the practical, and the practical became the obsolete”.

Jon Stewart

6.1 Discusión

6.1.1 El criterio previsor

Uno de los temas más importantes por tratar en este punto del documento es justamente responder la pregunta que se planteó en la figura 1.1: ¿es posible una red móvil descentralizada?

La respuesta del autor es sí. En el capítulo 2 el lector pudo conocer otras propuestas con grandes avances en esta búsqueda. De esta manera, el criterio previsor puede ser un gran complemento que ayude a manejar el reto de una red P2P móvil.

Con base en la evidencia estadística colectada y procesada en el capítulo 5 es posible concluir que:

Con el uso de un criterio previsor para la selección de pares en redes P2P móviles **medianamente densas** se obtiene una menor cantidad de desconexiones que al utilizar los criterios de señal, tiempo o azar.

6.1.2 Selección de los factores

La selección de factores es un paso muy importante al realizar un análisis de varianza sobre los resultados de un experimento. En [4] [37] hay secciones dedicadas a esta tarea.

Es fácil caer en la trampa y creer que cuantos más factores se consideren, más reveladores serán los resultados que arroje ANOVA.

No obstante, siempre es necesario recordar que los factores deben tener relación con la variable de respuesta, por lo que su escogencia debe tener algún fundamento, como las observaciones iniciales o las fuentes bibliográficas.

¿Cómo explicaría el efecto de los movimientos telúricos en esta investigación? Dichosamente, este no fue uno de los factores seleccionados, pues no se encontró referencia alguna en las fuentes investigadas.

Por otro lado, al modelar y ejecutar las simulaciones se pudo observar que el alcance de las antenas sí generó cambios en los resultados. Aun así, se decidió mantenerlo fijo, pues las limitantes de tiempo y alcance no permitieron su debido estudio para así proponer los niveles por considerar.

6.1.3 Redes densas

Para sorpresa del autor, durante esta investigación los resultados arrojados en la sección de análisis de varianza para las combinaciones de factores que incluían el tamaño de la red demostraron que las redes muy densas tienen un efecto bastante errático en cuanto a las desconexiones (figuras 5.10, 5.12, 5.13, 5.16, 5.19 y 5.22).

Si bien se esperaba que la cantidad de desconexiones creciera conforme la red se volviera más densa, el análisis ayudó a encontrar comportamientos como el descrito en la figura 5.14.

Lo anterior abre las puertas a una investigación centrada en este factor, con mayor rango de niveles. Tanto en la propuesta como la ejecución de esta investigación se utilizaron tres niveles, pues por ser la primera exploración con este enfoque se deseaba tener una visión macro de los resultados para establecer la factibilidad del criterio previsor.

En otras palabras, la búsqueda principal se centró aceptar o rechazar la hipótesis, pero estos resultados sirven como base para establecer que la densidad de la red es un camino de investigación promisorio.

6.2 Trabajo futuro

Tal como se mencionó en el capítulo de Introducción, la selección de pares es uno de los componentes de una red P2P. Por lo tanto, queda mucho trabajo por delante en la búsqueda de una red de telefonía descentralizada:

- Investigar el impacto de la densidad de una red distribuida en su desempeño: se propone ampliar la gama de niveles para establecer los puntos de convergencia o de ruptura.
- Caracterizar la ocurrencia de huérfanos dentro de la red: esto podría pasar no solo a nivel de agentes independientes, sino de grupos de agentes que pueden quedar totalmente descomunicados de otros (cliques).
- Búsqueda de nuevos factores: el hecho de que la evidencia observacional sugiera que el rango de la antena puede tener un efecto en la cantidad de desconexiones inclina al autor a pensar que la topografía también puede tener un efecto digno de consideración.
- Nuevos patrones de migración: como caso específico del punto anterior, se encontró que los patrones de migración en seres vivos es un campo de estudio activo y con amplios efectos en el estudio de la vida silvestre y simulaciones de carácter científico (como la presente investigación) o comercial (videojuegos). Incorporar los nuevos hallazgos del campo, y hasta generarlos a partir de este hilo de investigación, es una empresa digna de emprender.

6.3 Conclusiones

La investigación plasmada en este documento deja como resultados:

1. La evidencia estadística muestra que una red P2P que utilice un algoritmo de enrutamiento basado en un criterio previsor genera menos desconexiones que al emplear los criterios de redes convencionales.
2. Es posible mejorar el desempeño al reducir la cantidad de desconexiones, como resultado del uso de un criterio previsor para enrutamiento.
3. El algoritmo previsor permite desarrollar una red P2P cuyo tiempo de despliegue sea mínimo, pues no necesita de un proceso de entrenamiento para operar.
4. El algoritmo previsor tiene un costo computacional lineal ($\Theta(n)$).
5. La movilidad de los nodos en una red P2P desempeña un rol muy importante la estabilidad de una red P2P.
6. El algoritmo previsor puede ser un complemento de las tecnologías P2P que buscan descentralizar el sistema de telefonía para así ofrecer un mejor servicio.

Creación de corridas aleatorias

Listing A.1: generateTest.sh

```
#!/bin/bash

criteria="SG_TM_RA_FS"
steps="1000_5000_10000_15000"
size="1_2_3"
migration="CD_SQ_CO_RA"
replicas=20
maps_gen="./generate_scenario.sh"
simulator="/home/ameseguer/TEC/MAESTRIA/Tesis/build-RVO-Desktop-Debug/RVO"
maps_path="scenarios"
results_path="results"
maps_file="create_scenarios"
exe_file="run_scenarios"
results_file="results/index.txt"
id=0;
```

```

echo -n "" > $maps_file
echo -n "" > $exe_file
echo -n "" > $results_file

for c in $criteria ; do
  for s in $steps ; do
    for z in $size ; do
      for m in $migration; do
        for r in `seq $replicas`; do
          testID=`printf "%03d\n" $id`
          echo "$maps_gen_␣-m_␣$m_␣-s_␣$z_␣>␣$maps_path/scene$testID.xml" >>
          echo "$simulator_␣$maps_path/scene$testID.xml_␣$c_␣$s_␣>␣$results_
          echo "$testID_␣$c_␣$m_␣$z_␣$r_␣$s" >> $results_file
          id=$(( $id + 1 ))
        done
      done
    done
  done
done

chmod +x $exe_file
chmod +x $maps_file

```

Creación de escenarios aleatorios

Listing B.1: generateScenario.sh

```
#!/bin/bash

MODEL=RA          # CD=converge/diver  RA=Random  SQ=Squad
SIZE=1
MAX_Y=100
MAX_X=100
MAX_N=1
R_VALUE=1
RADIO=0.0001

while getopts ":m:s:" opt; do
    case $opt in
        m)
            MODEL=$OPTARG
            ;;
        s)
            ;;
    esac
done
```

```

        SIZE=$OPTARG
        ;;
    \?)
        echo "Invalid option: -$OPTARG" >&2
        exit 1
        ;;
    :)
        echo "Option -$OPTARG requires an argument." >&2
        exit 1
        ;;
esac
done

if [ "$SIZE" == "1" ] ; then
    MAX_WP=100
    MIN_WP=2
    MAX_AG=100
    MIN_AG=0
fi

if [ "$SIZE" == "2" ] ; then
    MAX_WP=100
    MIN_WP=2
    MAX_AG=800
    MIN_AG=100
fi

if [ "$SIZE" == "3" ] ; then
    MAX_WP=100
    MIN_WP=2
    MAX_AG=2000
    MIN_AG=800
fi

```

```

case $MODEL in

    "CD")
        R_VALUE=1;;
    "RA")
        R_VALUE=5;;
    "SQ")
        R_VALUE=5;
        SQ_NUMBER=$AG_NUMBER;
        AG_NUMBER=1;;
    "CO")
        R_VALUE=1;

esac

WP_NUMBER='echo $((( $MIN_WP+$RANDOM)%$MAX_WP))'
AG_NUMBER='echo $((( $MIN_AG+$RANDOM)%$MAX_AG))'

if [ $WP_NUMBER -lt $MIN_WP ] ; then
    WP_NUMBER=$MIN_WP
fi

if [ $AG_NUMBER -lt $MIN_AG ] ; then
    AG_NUMBER=$MIN_AG
fi

count=0

```

```

echo '<scenario type="'$MODEL'">'

##### WAYPOINTS

echo "<!--$WP_NUMBER waypoints define before the agents!-->"

case $MODEL in

    "CD")

        echo '<waypoint id="0" x="0" y="0" r="1" />'
        count=1 ;;

    esac

while [ $count -lt $WP_NUMBER ] ; do
    x_value='echo $(((($RANDOM+1)%($MAX_X*2))-MAX_X))'
    y_value='echo $(((($RANDOM+1)%($MAX_Y*2))-MAX_Y))'

    echo '<waypoint id="'$count'" x="'$x_value'" y="'$y_value'" r="'$R_VALUE'"

    count=$((count+1))
done

##### AGENTS

count=0

```

```

echo "<!--_␣$AG_NUMBER_␣agents_-->"

while [ $count -lt $AG_NUMBER ] ; do

    x_value='echo $((((($RANDOM+1)%($MAX_X*2))-MAX_X))'
    y_value='echo $((((($RANDOM+1)%($MAX_Y*2))-MAX_Y))'
    w_total='echo $((((($RANDOM+1)%WP_NUMBER))' #random number of waypoints
    w_count=0

    if [ "$MODEL" != "SQ" ] ; then
        echo '<agent id="'$count' " x="'$x_value' " y="'$y_value' " n="1" dx=
    else
        echo '<agent id="'$count' " x="'$x_value' " y="'$y_value' " n="'$SQ_N
    fi

    case $MODEL in

        "CD")

            echo '    <addwaypoint id="0" />'
            w_value='echo $((((($RANDOM+1)%WP_NUMBER))'
            echo '    <addwaypoint id="'$w_value' " />' ;;

        "CO")

            while [ $w_count -lt 2 ] ; do

                w_value='echo $((((($RANDOM+1)%WP_NUMBER))'
                echo '    <addwaypoint id="'$w_value' " />'
                w_count=$((w_count+1))

            done ;;

        "RA" | "SQ")

            while [ $w_count -lt $w_total ] ; do

                w_value='echo $((((($RANDOM+1)%WP_NUMBER))'
                echo '    <addwaypoint id="'$w_value' " />'

```

```
        w_count=$((w_count+1))
    done ;;
esac

echo "</agent>"

count=$((count+1))
done

echo "</scenario>"
```

Ejemplo de un escenario CD

```
<scenario type="CD">
  <!-- 2 waypoints - define before the agents! -->
  <waypoint id="0" x="0" y="0" r="1" />
  <waypoint id="1" x="8" y="61" r="1" />
  <!-- 10 agents -->
  <agent id="0" x="77" y="51" n="1" dx="70" dy="10" r="0.0001">
    <addwaypoint id="0" />
    <addwaypoint id="0" />
  </agent>
  <agent id="1" x="-37" y="85" n="1" dx="70" dy="10" r="0.0001">
    <addwaypoint id="0" />
    <addwaypoint id="0" />
  </agent>
  <agent id="2" x="82" y="71" n="1" dx="70" dy="10" r="0.0001">
    <addwaypoint id="0" />
    <addwaypoint id="1" />
  </agent>
</scenario>
```

```
</agent>
<agent id="3" x="-88" y="-90" n="1" dx="70" dy="10" r="0.0001">
  <addwaypoint id="0" />
  <addwaypoint id="1" />
</agent>
<agent id="4" x="50" y="46" n="1" dx="70" dy="10" r="0.0001">
  <addwaypoint id="0" />
  <addwaypoint id="1" />
</agent>
<agent id="5" x="72" y="-11" n="1" dx="70" dy="10" r="0.0001">
  <addwaypoint id="0" />
  <addwaypoint id="1" />
</agent>
<agent id="6" x="48" y="-64" n="1" dx="70" dy="10" r="0.0001">
  <addwaypoint id="0" />
  <addwaypoint id="1" />
</agent>
<agent id="7" x="19" y="-53" n="1" dx="70" dy="10" r="0.0001">
  <addwaypoint id="0" />
  <addwaypoint id="1" />
</agent>
<agent id="8" x="-67" y="-38" n="1" dx="70" dy="10" r="0.0001">
  <addwaypoint id="0" />
  <addwaypoint id="0" />
</agent>
<agent id="9" x="-84" y="11" n="1" dx="70" dy="10" r="0.0001">
  <addwaypoint id="0" />
  <addwaypoint id="0" />
</agent>
</scenario>
```

Acerca de PEDSIM

PEDSIM [25] es un sistema microscópico de simulación de multitudes de peatones. Es adecuado para su uso en simulaciones de multitudes (por ejemplo, la simulación de interior de evacuación, simulaciones a gran escala al aire libre y, por supuesto, la de usuarios de teléfonos móviles), donde uno está interesado en una salida como la densidad de peatones o el tiempo de evacuación. Además, la calidad de la trayectoria del agente individual es lo suficientemente alta para crear animaciones multitud de peatones masivas (por ejemplo, para las películas o la visualización arquitectónica).

La biblioteca *libpedsim* es fácil de usar y extender; es un buen punto de partida para proyectos de ciencia, por ejemplo, imágenes, películas cortas, capturas de pantalla y, claro está, redes P2P.

Asimismo, permite utilizar la dinámica de peatones en su propio *software*. Sobre la base de C++ puro y sin paquetes de *software* adicionales, funciona prácticamente en cualquier sistema operativo.

Los peatones son visibles en la interfaz de usuario en tiempo real. Con el uso del archivo o una salida de conexión TCP, el procesamiento por lotes o en tiempo real es

posible. Típicamente, la salida de PEDSIM se alimenta de un motor de representación gráfica donde se crean los seres humanos con aspecto realista; estos caminan sobre la base de las trayectorias generadas por PEDSIM.

Por su parte, el modelo de simulación utilizado es simple y fácil de entender; además, es posible extender el comportamiento del agente. Varias de estas estrategias mentales se incluyen, y es posible al añadir su propio código.

PEDSIM tiene tres componentes principales:

- **La escena:** el espacio donde los agentes van a movilizarse.
- **Los puntos de ruta:** coordenadas cartesianas que representan una posición dentro de la escena.
- **Los agentes:** con base en una serie de puntos de ruta, siguen una trayectoria dentro de la escena. La inteligencia artificial básica es provista con la biblioteca (colisión con otros agentes, congestión, obstáculos, etc).

Muestra de los resultados del simulador

ID	ALGORITHM	MIGRATION	SIZE	REPLICA	STEPS	DISCONNECTS
893	SG	CD	3	14	15000	639233
3030	FS	RA	2	11	1000	3505
1676	TM	RA	3	17	10000	84701
3810	FS	CD	3	11	15000	78879
745	SG	SQ	1	6	15000	0
3525	FS	CD	3	6	10000	54476
3549	FS	SQ	3	10	10000	20720
925	SG	CD	3	6	15000	934387
1728	TM	CD	1	9	15000	0
780	SG	RA	1	1	15000	0
3069	FS	SQ	3	10	1000	2416
3699	FS	CD	2	20	15000	43146
1296	TM	CD	2	17	5000	10042
3258	FS	CD	2	19	5000	22595
2345	RA	SQ	3	6	5000	16241

Bibliografía

- [1] Sarita V. Adve and Kourosh Gharachorloo. Shared memory consistency models: A tutorial. *Computer*, 29(12):66–76, December 1996. ISSN 0018-9162. doi: 10.1109/2.546611. URL <http://dx.doi.org/10.1109/2.546611>.
- [2] M. Afergan. Using repeated games to design incentive-based routing systems. In *INFOCOM 2006. 25th IEEE International Conference on Computer Communications. Proceedings*, pages 1–13, 2006. doi: 10.1109/INFOCOM.2006.61.
- [3] A. V. Aho, J. E. Hopcroft, and J. D. Ullman. *The Design and Analysis of Computer Algorithms*. Addison-Wesley Publishing Company, 1974.
- [4] M.J. Anderson and P.J. Whitcomb. *DOE Simplified: Practical Tools for Effective Experimentation, Third Edition*. CRC Press, 2015. ISBN 9781498730907. URL https://books.google.co.cr/books?id=_TKsCQAAQBAJ.
- [5] Capt. Diana Parzik U.S. Army. Una casa derruída tras el terremoto, 2009. URL <http://commons.wikimedia.org/wiki/File:2009CostaRicaEarthquake2.jpg>.
- [6] Eric Weisstein at Wolfram Research. Gompertz curve, 2014. URL <http://mathworld.wolfram.com/GompertzCurve.html>. [Online; accessed 15-October-2014].
- [7] Forouzan B.A and Chung Fegan S. *Data Communication and Networking, Fourth Edition*. McGraw Hill, 2007. ISBN 0-07-296775-7.
- [8] Paul E. Black. Euclidean distance, 2004. URL <http://www.nist.gov/dads/HTML/euclidndstnc.html>. [Online; accessed 15-October-2014].
- [9] Stevens Le Blond, Fabrice Le Fessant, and Erwan Le Merrer. Choosing partners based on availability in p2p networks. *ACM Trans. Auton. Adapt. Syst.*, 7(2): 25:1–25:14, July 2012. ISSN 1556-4665. doi: 10.1145/2240166.2240175. URL <http://doi.acm.org/10.1145/2240166.2240175>.

- [10] Box, G. E. P. and Cox, D. R. An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 26(2):211–252, 1964. ISSN 0035-9246.
- [11] Arne Buthmann. Making data normal using box-cox power transformation, Unknown. URL <http://www.isixsigma.com/tools-templates/normality/making-data-normal-using-box-cox-power-transformation/>. [Online; accessed 15-October-2014].
- [12] Michael Cohrs. Implementation and evaluation of the peer-to-peer-protocol (p2pp) for p2psip, March 2008. ISSN 1612-6793. URL <http://www.net.informatik.uni-goettingen.de/publications/1528/thesis-mcohrs.pdf>.
- [13] Richard I. Cook. How complex systems fail, 2000.
- [14] G. Coulouris, J. Dollimore, T. Kindberg, and G. Blair. *Distributed Systems: Concepts and Design*. Pearson Education, 2011. ISBN 9780132143011.
- [15] Oficina del informe sobre desarrolllo humano. Mapa, 2010. URL <http://hdr.undp.org/es/datos/mapa/>. [Online; accessed 25-May-2013].
- [16] La Oficina del Informe sobre Desarrollo Humano. Mapa, indicadores internacionales sobre desarrollo humano, 2013. URL <http://hdr.undp.org/es/datos/mapa/>.
- [17] Somayeh Dodge, Robert Weibel, and Anna-Katharina Lautenschütz. Towards a taxonomy of movement patterns. *Information Visualization*, 7(3):240–252, June 2008. ISSN 1473-8716. doi: 10.1057/palgrave.ivs.9500182. URL <http://dx.doi.org/10.1057/palgrave.ivs.9500182>.
- [18] Michel Dubois, Christoph Scheurich, and Faye Briggs. Memory access buffering in multiprocessors. In *25 years of the international symposia on Computer architecture (selected papers)*, ISCA '98, pages 320–328, New York, NY, USA, 1998. ACM. ISBN 1-58113-058-9. doi: 10.1145/285930.285991. URL <http://doi.acm.org/10.1145/285930.285991>.
- [19] Frank H. P. Fitzek and Hassan Charaf. *Mobile Peer-to-Peer Networks: An Introduction to the Tutorial Guide*, pages 1–18. John Wiley Sons, Ltd, 2009. ISBN 9780470747889. doi: 10.1002/9780470747889.ch1. URL <http://dx.doi.org/10.1002/9780470747889.ch1>.
- [20] Fon.com. What is fon, 2013. URL <http://corp.fon.com/#what-is-fon>. [Online; accessed 25-May-2013].
- [21] Leonardo Arias Fonseca. Caracterización e inducción de comportamientos emergentes en redes de pares. Master’s thesis, Instituto Tecnológico de Costa Rica, 2013.

- [22] Bertalan Forstner and Imre Kelényi. *Mobile Social Networking – Beyond the Hype*, pages 161–190. John Wiley Sons, Ltd, 2009. ISBN 9780470747889. doi: 10.1002/9780470747889.ch9. URL <http://dx.doi.org/10.1002/9780470747889.ch9>.
- [23] Free Software foundation. Gnu bash, 2014. URL <http://www.gnu.org/software/bash/>. [Online; accessed 15-March-2014].
- [24] Seth Gilbert and Nancy Lynch. Brewer’s conjecture and the feasibility of consistent, available, partition-tolerant web services. *SIGACT News*, 33(2):51–59, June 2002. ISSN 0163-5700. doi: 10.1145/564585.564601. URL <http://doi.acm.org/10.1145/564585.564601>.
- [25] Christian Gloor. Pedsim — a microscopic pedestrian crowd simulation system, 2014. URL <http://pedsim.silmaril.org/>. [Online; accessed 15-March-2014].
- [26] GNU. Regarding gnutella, 2013. URL <http://www.gnu.org/philosophy/gnutella.html>. [Online; accessed 25-May-2013].
- [27] Philippe Golle, Kevin Leyton-Brown, Ilya Mironov, and Mark Lillibridge. Incentives for sharing in peer-to-peer networks. In *Proceedings of the Second International Workshop on Electronic Commerce, WELCOM '01*, page 75–87, London, UK, 2001. Springer-Verlag, Springer-Verlag. ISBN 3-540-42878-X. URL <http://portal.acm.org/citation.cfm?id=647064.714876>.
- [28] Jim Gray. Why do computers stop and what can be done about it?, 1985.
- [29] The World Bank Group. World development indicators:power and communications, 2014. URL <http://data.worldbank.org/data-catalog/world-development-indicators>. [Online; accessed 15-October-2014].
- [30] Federico Herrera. Cinchona, 2009. URL <http://recordandocinchona.blogspot.com>. [Online; accessed 25-May-2013].
- [31] Jabber.org. Jabber.org faq, 2013. URL <http://www.jabber.org/faq.html>. [Online; accessed 25-May-2013].
- [32] W. Kellerer. Dienstarchitekturen in der telekommunikation-evolution, methoden und vergleich. Technical report, Technical Report TUM-LKN-TR-9801, 1998.
- [33] L. Lamport. How to make a multiprocessor computer that correctly executes multiprocess programs. *Computers, IEEE Transactions on*, C-28(9):690–691, 1979. ISSN 0018-9340. doi: 10.1109/TC.1979.1675439.
- [34] S.I. Lee, Hyunggon Park, and M. van der Schaar. Foresighted joint resource reciprocation and scheduling strategies for real-time video streaming over peer-to-peer networks. In *Packet Video Workshop, 2009. PV 2009. 17th International*, pages 1–10, 2009. doi: 10.1109/PACKET.2009.5152153.

- [35] Lisa M Lix, Joanne C Keselman, and HJ Keselman. Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance f test. *Review of Educational Research*, 1996.
- [36] Haim Matzner and Kirk T. McDonald. Isotropic radiators. 2003.
- [37] Douglas C. Montgomery. *Design and Analysis of Experiments*. John Wiley and Sons, 2001. ISBN 0471316490.
- [38] Mary Natrella. *NIST/SEMATECH e-Handbook of Statistical Methods*. NIST/SEMATECH, July 2010. URL <http://www.itl.nist.gov/div898/handbook/>.
- [39] University of North Carolina at Chapel Hill. Rvo2 library: Reciprocal collision avoidance for real-time multi-agent simulation, 2014. URL <http://gamma.cs.unc.edu/RV02/>. [Online; accessed 1-November-2015].
- [40] Ooma®. Ooma homepage, 2013. URL <http://www.ooma.com/how-ooma-works>. [Online; accessed 25-May-2013].
- [41] H. Pham and Chin-Diew Lai. On recent generalizations of the weibull distribution. *Reliability, IEEE Transactions on*, 56(3):454–458, Sept 2007. ISSN 0018-9529. doi: 10.1109/TR.2007.903352.
- [42] OpenVoip Project. Openvoip project, 2009. URL <http://www1.cs.columbia.edu/~salman/peer/>. [Online; accessed 25-May-2013].
- [43] R Project. What is r, 2013. URL <http://www.r-project.org/about.html>. [Online; accessed 25-May-2013].
- [44] Mercedes Agüero R. Ice enfrenta pico de demanda celular en medio de averías, 2010. URL <http://www.nacion.com/2010-12-23/ElPais/NotasSecundarias/ElPais2630803.aspx?Page=2>. [Online; accessed 25-May-2013].
- [45] R. Rao and S. Vrudhula. Battery optimization vs energy optimization: Which to choose and when? In *Proceedings of the 2005 IEEE/ACM International Conference on Computer-aided Design, ICCAD '05*, pages 439–445, Washington, DC, USA, 2005. IEEE Computer Society. ISBN 0-7803-9254-X. URL <http://dl.acm.org/citation.cfm?id=1129601.1129664>.
- [46] Giovanni Sanabria Brenes. *Comprendiendo la estadística inferencial*. Editorial Tecnológica de Costa Rica, 2011. ISBN 78-9977-66-238-1.
- [47] Mónica Cordero Sancho. Deaths from earthquakes in 2012, 2012. URL http://www.elfinancierocr.com/tecnologia/Sismo-celular_0_148785122.html. [Online; accessed 25-May-2013].
- [48] Nicola Santoro. *Design and Analysis of Distributed Algorithms (Wiley Series on Parallel and Distributed Computing)*. Wiley-Interscience, 2006. ISBN 0471719978.

- [49] Eric Savitz. The crucial difference between wireless and mobile. 2012. URL <http://www.forbes.com/sites/ciocentral/2012/02/21/the-crucial-difference-between-wireless-and-mobile/>.
- [50] R. Schollmeier. A definition of peer-to-peer networking for the classification of peer-to-peer architectures and applications. In *Peer-to-Peer Computing, 2001. Proceedings. First International Conference on*, pages 101–102, 2001. doi: 10.1109/P2P.2001.990434.
- [51] Clay Shirky. What is p2p... and what isn't? (june 21st, 2011), 2000. URL <http://openp2p.com/pub/a/p2p/2000/11/24/shirky1-whatisp2p.html>.
- [52] Abraham Silberschatz, Peter Baer Galvin, and Greg Gagne. *Operating System Concepts*. John Wiley & Sons, Inc., New York, NY, USA, 7th edition, 2004. ISBN 0471694665.
- [53] William Stallings. *Wireless Communications and Networks*. Prentice Hall, 2005. ISBN 0130408646.
- [54] Andrew S. Tanenbaum and David J. Wetherall. *Computer Networks*. Prentice Hall Press, Upper Saddle River, NJ, USA, 5th edition, 2010. ISBN 0132126958, 9780132126953.
- [55] TerraNet®. Mesh, 2013. URL <http://terranet.se/about/technology/mesh-technology>. [Online; accessed 25-May-2013].
- [56] Narseo Vallina-Rodriguez, Christos Efstratiou, Geoffrey Xie, and Jon Crowcroft. Enabling opportunistic resources sharing on mobile operating systems: Benefits and challenges. In *Proceedings of the 3rd ACM Workshop on Wireless of the Students, by the Students, for the Students, S3 '11*, pages 29–32, New York, NY, USA, 2011. ACM. ISBN 978-1-4503-0868-7. doi: 10.1145/2030686.2030696. URL <http://doi.acm.org/10.1145/2030686.2030696>.
- [57] W3C®. Extensible markup language (xml), 2014. URL <http://www.w3.org/XML/>. [Online; accessed 15-March-2014].
- [58] R.E. Walpole, R.H. Myers, and S.L. Myers. *Probability and Statistics for Engineers and Scientists*. Pearson Education, Limited, 2010. ISBN 9780321629111. URL <http://books.google.co.cr/books?id=tzZxRQAACAAJ>.
- [59] Wikipedia. Bittorrent (protocol) — wikipedia, the free encyclopedia, 2012. URL [http://en.wikipedia.org/w/index.php?title=BitTorrent_\(protocol\)&oldid=479301802](http://en.wikipedia.org/w/index.php?title=BitTorrent_(protocol)&oldid=479301802). [Online; accessed 7-March-2012].
- [60] Wikipedia. Napster — wikipedia, the free encyclopedia, 2012. URL <http://en.wikipedia.org/w/index.php?title=Napster&oldid=478961974>. [Online; accessed 25-May-2013].

- [61] Li Xi, Ji Hong, F.R. Yu, and Zheng Ruiming. A fcm-based peer grouping scheme for node failure recovery in wireless p2p file sharing. In *Communications, 2009. ICC '09. IEEE International Conference on*, pages 1–5, 2009. doi: 10.1109/ICC.2009.5198602.

F.1 Creative Commons Atribución – Compartir Igual 3.0 Costa Rica

La Obra se otorga bajo los términos del presente contrato de Licencia, está protegida por las leyes y los instrumentos internacionales debidamente suscritos en materia de derechos de autor y derechos conexos y los reglamentos que las desarrollan. Queda prohibido cualquier uso distinto del autorizado bajo la presente licencia o la legislación de derechos de autor y derechos conexos aplicables.

Al ejercer cualquiera de los derechos aquí otorgados, usted acepta y acuerda quedar obligado por los términos del presente contrato de licencia, en la medida en que esto se considere un contrato de acuerdo a la ley aplicable. El Licenciatario le concede los derechos aquí contenidos en consideración de la aceptación de tales términos y condiciones.

1. Definiciones

“Autor Original” significa, en el caso de una obra literaria, científica o artística, el individuo, personas, o entidades que crearon la Obra, o si ninguna persona o entidad puede ser identificada, el editor, y además (i) en el caso de interpretaciones: los actores, cantantes, músicos, bailarines y otras personas que representen, canten, reciten, declamen, interpreten o ejecuten en cualquier forma obras literarias o artísticas o expresiones del folclore, (ii) en el caso de fonogramas: el productor es la persona física o jurídica que fija por primera vez los sonidos de una ejecución u otros sonidos, y, (iii) en el caso de las emisiones, la organización que las transmite. “Colección”: es la Obra constituida por un conglomerado de obras literarias, científicas o artísticas, tales como enciclopedias y antologías, interpretaciones, ejecuciones, fonogramas, emisiones y transmisiones. Estas Obras son incluidas en la Colección en su totalidad y de forma inalterada junto con una o más contribuciones, constituyendo cada una obras separadas e independientes en sí mismas, que en conjunto se integran en un todo colectivo. Una Colección no será considerada una Obra Derivada para los fines de esta Licencia. “Distribuir” significa poner a disposición del público el original y copias de la Obra u Obra Derivada, según corresponda, mediante cualquier título traslativo de dominio,

oneroso o gratuito, transitorio o permanente. “Ejecución Pública” se refiere a llevar a cabo la interpretación en público de la Obra y comunicar al público tales interpretaciones por cualquier medio o procedimiento, incluyendo medios alámbricos, inalámbricos o digitales; poner a disposición del público Obras de tal forma que los miembros del público puedan acceder a estas Obras desde el lugar que ellos elijan; asimismo interpretar la obra al público por cualquier medio o procedimiento con el fin de transmitir y retransmitir la obra por cualquier medio, incluso los signos, sonidos o imágenes. “Elementos de la Licencia” significan los siguientes atributos de licencia seleccionados por el Licenciante e indicados en el título de esta Licencia: Atribución, CompartirIgual. “Licencia Compatible de Creative Commons” significa una licencia enumerada en <http://creativecommons.org/compatible> que ha sido aprobada por Creative Commons esencialmente como un equivalente a esta Licencia, incluyendo, como mínimo, una licencia que: (i) contiene términos que tienen el mismo propósito, el significado y efecto que los Elementos de Licencia de esta Licencia, y (ii) permite expresamente la redistribución de obras derivadas puestas a disposición por el licenciante en virtud de esta Licencia o una licencia de Creative Commons de otras jurisdicciones que contengan los mismos Elementos de Licencia que esta Licencia. “Licenciante” significa el individuo, las personas, entidad o entidades que ofrecen la Obra bajo los términos de esta Licencia. “Obra” significa el trabajo literario, científico o artístico que se ofrece bajo los términos de esta Licencia, cualquier producción en el campo literario, científico y artístico, cualquiera que sea el modo o forma de expresión, incluidos los formatos digitales. Esto incluye, sin limitación: folletos, panfletos, conferencias, presentaciones, sermones u otras obras de la misma naturaleza; obras dramáticas o dramático-musicales; obras coreográficas o de entretenimiento en pantomimas; composiciones musicales con letra o sin ella; obras cinematográficas u otras expresadas por procedimiento análogo a la cinematografía; obras de dibujo, pintura, arquitectura, escultura, grabado, litografía; trabajo fotográficos o trabajos expresados por procedimiento análogo a la fotografía; mapas, planos, croquis u obras tridimensionales en relación con la geografía, la topografía, la arquitectura o las ciencias; espectáculos; emisiones; fonograma; recopilaciones de datos en la medida en que estén protegidas como derechos de autor; trabajos realizados para un espectáculo de variedad o un artista de circo en la medida en que sean considerados una obra literaria o artística. “Obra Derivada” significa una obra basada en o sobre la Obra y otras obras preexistentes, tales como traducciones, adaptaciones, obras derivadas, arreglos musicales u otras alteraciones de una obra literaria, científica, artística, musical, fonograma o ejecución. Esto incluye también adaptaciones cinematográficas o cualquier otra forma en la cual la Obra puede ser reformulada, transformada o adaptada, incluyendo cualquier forma reconocible derivada del original. A manera de clarificación, cuando la Obra sea una obra musical, presentación, o fonograma, la sincronización de la Obra con una imagen móvil (“synching”) será considerada una Obra Derivada para el propósito de esta Licencia. “Reproducir” significa el hacer copias de la Obra por cualquier medio incluyendo sin limitación, las grabaciones sonoras o visuales, y el derecho de fijación y reproducción de las fijaciones de la Obra, incluido el almacenamiento de una interpretación o ejecución

protegida o de un fonograma en forma digital o cualquier otro medio electrónico. “Usted” significa un individuo o entidad ejerciendo los derechos bajo esta Licencia, quien previamente no ha contravenido los términos de esta Licencia con respecto a la Obra, o que haya recibido el permiso expreso del Licenciante para ejercer derechos bajo esta Licencia pese a haber ejercido una contravención anterior.

2. Excepciones. Ninguna disposición presente en esta Licencia tiene por objeto reducir, limitar, o restringir derechos existentes, ni afectar las excepciones presentes en la Ley de Derechos de Autor y Derechos Conexos ni en otras leyes aplicables.
3. Otorgamiento de Licencia. Sujeto a los términos y condiciones de esta Licencia, el Licenciante otorga a Usted una licencia mundial, libre de regalías, no exclusiva, perpetua (por la duración de los derechos de autor aplicables) para ejercer los derechos en la Obra que se indican a continuación:

Reproducir la Obra, incorporar la Obra en una o más Colecciones, y Reproducir la Obra incorporada en las Colecciones. Crear y Reproducir Obras Derivadas, bajo la condición de que dicha Obra Derivada tome las medidas razonables para identificar, marcar y/o demarcar claramente de otra forma los cambios que se hicieron a la Obra original. Por ejemplo, una traducción puede marcarse como “La obra original fue traducida del Inglés al Español”, o una modificación podría indicar “La obra original ha sido modificada.” Distribuir y Ejecutar Públicamente la Obra, incluyendo las incorporadas en Colecciones. Distribuir y Ejecutar Públicamente Obras Derivadas. Para evitar cualquier duda: Sistemas de licencias obligatorias no renunciables. En donde no sea posible renunciar al derecho a cobrar regalías a través de cualquier sistema de licencias estatutarias u obligatorias, el Licenciante se reserva el derecho exclusivo a cobrar derechos de autor para cualquier ejercicio que Usted haga de los derechos concedidos bajo esta Licencia. Sistemas de licencias obligatorias renunciables. En donde el derecho a cobrar regalías a través de cualquier sistema de licencias estatutarias u obligatorias es renunciable, el Licenciante se reserva el derecho exclusivo a cobrar derechos de autor para cualquier ejercicio que Usted haga de los derechos concedidos bajo esta licencia. Sistemas de licencias voluntarias. El Licenciante se reserva el derecho a cobrar regalías por el ejercicio hecho por Usted de cualquiera de los derechos otorgados por medio de esta licencia, ya sea individualmente, o si el Licenciante sea un miembro de una sociedad de gestión colectiva que administre los regímenes de licencias no voluntarias, por medio de tal sociedad.

Los derechos mencionados anteriormente pueden ser ejercidos en todos los medios y formatos ahora conocidos o por inventarse. Los derechos antes mencionados incluyen el derecho a efectuar las modificaciones que sean técnicamente necesarias para ejercer los derechos en otros medios y formatos. Todos los derechos no otorgados expresamente por el Licenciante se encuentran reservados.

4. Restricciones. La licencia otorgada en la anterior Sección 3 está expresamente sujeta y limitada por las siguientes restricciones:

Usted puede Distribuir o Ejecutar Públicamente la Obra sólo bajo los términos de esta Licencia. Usted debe incluir una copia de esta Licencia, o en su defecto, un vínculo con cada copia de la Obra que Distribuya o Ejecute Públicamente. Usted no podrá ofrecer o imponer ninguna condición sobre la Obra que restrinja los términos de esta Licencia, o la capacidad del usuario de la Obra para ejercer los derechos otorgados al usuario bajo los términos de la Licencia. Usted no puede sub-licenciar la Obra. Usted debe mantener intactos todos los avisos que se refieran a esta Licencia con cada copia de la Obra que Distribuya o Ejecute Públicamente. Cuando Usted Distribuya o Ejecute Públicamente la Obra, no puede imponer cualquier medida tecnológica efectiva que restrinja la capacidad de un destinatario de la Obra para ejercer los derechos otorgados al receptor bajo los términos de esta Licencia. Esta Sección 4 (a) se aplica a la Obra incorporada en una Colección, pero esto no exige que la Colección aparte de la Obra misma sea sujeta a los términos de esta Licencia. Si Usted crea una Colección, previo aviso de cualquier Licenciante, Usted debe, en la medida de lo posible, quitar de la misma cualquier crédito requerido en la Sección 4 (c), según lo solicitado. Si Usted crea una Obra Derivada, previo aviso de cualquier Licenciante, Usted debe, en la medida de lo posible, quitar de la misma cualquier crédito requerido en la Sección 4 (c), según lo solicitado. Usted puede Distribuir o Ejecutar Públicamente una Obra Derivada sólo bajo las condiciones de: (i) esta Licencia; (ii) una versión posterior a esta Licencia con los mismos Elementos de Licencia; (iii) una licencia de Creative Commons de otra jurisdicción (ya sea esta o una versión de licencia posterior) que contenga los mismos Elementos de Licencia, por ejemplo, la licencia Attribution-ShareAlike 3.0 Unported; (iv) una Licencia Compatible de Creative Commons. Si se distribuye la Obra Derivada bajo una de las licencias mencionadas en este punto (iv), debe cumplir con los términos de esa licencia. Si usted distribuye la Obra Derivada bajo los términos de cualquiera de las licencias mencionadas en (i), (ii) o (iii) (la “Licencia Aplicable”), debe cumplir con los términos generales de la Licencia Aplicable y con las siguientes disposiciones: (I) Usted debe incluir una copia de, o un vínculo a, esta Licencia en cada copia de cada Obra Derivada que Usted Distribuya o Ejecute Públicamente. (II) Usted no podrá ofrecer o imponer ninguna condición sobre la Obra Derivada que restrinja los términos de la Licencia Aplicable o la capacidad del usuario de la Obra Derivada al ejercicio de los derechos concedidos al usuario, según los términos de la Licencia Aplicable. (III) Usted debe mantener intactos todos los avisos que se refieren a la Licencia Aplicable con cada copia de la Obra tal como se prevé en la Obra Derivada que Usted Distribuya o Ejecute Públicamente; (IV) Cuando Usted Distribuya o Ejecute Públicamente la Obra Derivada, no puede imponer ninguna medida tecnológica efectiva que restrinja la capacidad de un destinatario de la Obra Derivada para ejercer los derechos otorgados al usuario bajo los términos de la Licencia Aplicable. La presente Sección 4 (b) aplica a la Obra Derivada cuando se incorpore en una Colección, pero esto no exige que la Colección, aparte de la Obra Derivada en sí, esté sujeta a los términos de esta Licencia. Si Usted Distribuye o Ejecuta Públicamente la Obra o cualquier Obra Derivada o Colección, debe, a menos que una solicitud haya sido hecha de

conformidad con la Sección 4 (a), mantener intactos todos los avisos de derechos de autor sobre la Obra. Asimismo debe proporcionar de manera razonable con respecto al medio utilizado: (I) el nombre del Autor Original (o seudónimo, en su caso) si fue suministrado, y/o si el Autor Original y/o Licenciante designara a otra parte (por ejemplo, institución patrocinadora, ente editor, revista, etc.) para la atribución (“Partes de Atribución”) en la notificación de derechos de autor del Licenciante, términos de servicio, o por otros medios razonables, el nombre de dicha Parte o Partes. (II) El título de la Obra si este se provee. (III) En la medida de lo razonablemente posible, un vínculo de Internet en caso que el Licenciante especifique que quiere ser asociado con la Obra, salvo que tal vínculo no se refiera a la notificación de derechos de autor o la información de Licencia. (IV) De conformidad con la Sección 3 (b), en el caso de una Obra Derivada, atribuir la autoría identificando el uso de la Obra en la Obra Derivada (por ejemplo, “Traducción Francesa de la Obra del Autor Original,” o “Guión cinematográfico basado en la Obra original del Autor Original”). Los créditos requeridos por esta Sección 4 (c) pueden ser implementados de cualquier forma razonable, a condición que en el caso de una Obra Derivada o Colección, dichos créditos aparezcan si un crédito para todos los autores que contribuyeron a la Obra Derivada o Colección aparece, a continuación, como parte de estos créditos y de una manera al menos tan destacada como los créditos de los demás autores que contribuyeron. Para evitar dudas, Usted sólo podrá utilizar el crédito requerido por esta Sección con el propósito de la atribución en la forma prevista anteriormente y, con el fin de ejercer sus derechos bajo esta Licencia. Usted no podrá implícita o explícitamente afirmar o implicar que existe conexión con el Autor Original, el Licenciante y/o Partes de Atribución, ni puede dar a entender que goza del patrocinio o respaldo de los mismos, a menos que haya obtenido de manera previa la expresa autorización por escrito por parte del Autor Original, Licenciante y/o Partes de Atribución. Los derechos morales de autor no se verán afectados en la presente Licencia en la medida que sean reconocidos y no puedan renunciarse de acuerdo a la ley aplicable.

5. Representaciones, Garantías y Responsabilidad

A menos que exista un acuerdo contrario por escrito entre las partes, y en la mayor medida permitida por la legislación aplicable, el Licenciante ofrece la Obra tal como se encuentra y no hace ninguna declaración o garantía de ningún tipo con respecto de la Obra, ya sea expresa, implícita, legal o de otro tipo, incluyendo sin limitación, garantías de título, comerciabilidad, idoneidad para un propósito particular, no infracción, o la ausencia de errores latentes u otros defectos, exactitud, o la presencia o ausencia de errores, sean o no sean descubiertos de forma previa.

6. Límites de responsabilidad. Excepto en lo requerido por la legislación aplicable, el Licenciante no asumirá ante Usted ninguna responsabilidad contractual o extracontractual por cualquier tipo de daños y perjuicios derivados del uso o mal funcionamiento, proveniente de esta Licencia o del uso de la obra, incluso si se ha notificado al

Licenciante de la posibilidad de tales daños.

7. Finalización

La presente Licencia y los derechos aquí concedidos finalizarán automáticamente en caso que Usted contravenga los términos de esta Licencia. Las personas o entidades que han recibido Obras Derivadas o Colecciones bajo esta Licencia, no tendrán terminadas sus Licencias siempre y cuando estos individuos o entidades sigan cumpliendo íntegramente con los términos aquella. Las secciones 1, 2, 5, 6, 7, y 8 subsistirán cualquier terminación de esta Licencia. Sujeto a los términos y condiciones antes mencionados, la licencia otorgada aquí es perpetua (mientras duren los derechos de autor de la obra). No obstante lo anterior, el Licenciante se reserva el derecho a publicar la Obra bajo diferentes términos, o de detener la distribución de la Obra en cualquier momento, siempre que esa elección no sirva para revocar los efectos de la presente Licencia (o cualquier otra licencia que haya sido concedida bajo los términos de esta Licencia). Esta Licencia continuará en pleno vigor y efecto a menos que termine como se ha indicado anteriormente.

8. Varios

Cada vez que Usted Distribuya o Ejecute Públicamente la Obra o una Colección, el Licenciante ofrece a los destinatarios una licencia en los mismos términos y condiciones que la Licencia concedida a Usted. Cada vez que Usted Distribuya o Ejecute Públicamente una Obra Derivada, el Licenciante ofrece a los destinatarios una licencia para la Obra original en los mismos términos y condiciones que la licencia concedida a Usted. Si alguna disposición de esta Licencia es inválida o inaplicable bajo la legislación aplicable, esto no afectará la validez o aplicabilidad del resto de los términos de esta Licencia y, sin mediar acción de las partes de este contrato, tal disposición será reformada en la mínima medida necesaria para que tal disposición sea válida y aplicable. Ningún término o disposición de esta Licencia se estimará renunciada a menos que medie consentimiento por escrito y firmado por las partes que serán afectadas por tal renuncia o consentimiento. Esta Licencia constituye el acuerdo completo entre las partes con respecto a la presente Obra. No existen otros entendimientos, acuerdos o representaciones con respecto a la Obra que no estén especificados aquí. El Licenciante no será obligado por ninguna disposición adicional que pueda aparecer en cualquier comunicación proveniente de Usted. Esta Licencia no puede ser modificada sin el acuerdo mutuo por escrito del Licenciante y Usted. Los derechos concedidos y la materia de referencia, en esta Licencia se elaboraron utilizando la terminología de la Ley 6683 sobre Derechos de Autor y Derechos Conexos, la Ley 7968 de Aprobación del Tratado de La OMPI sobre Derechos De Autor (WCT) (1996), la ley 8039 De Procedimientos de Observancia de Derechos de Propiedad Intelectual, el Reglamento a la Ley de Derechos de Autor y Derechos (N° 24611-J), la Convención de Berna para la Protección de Obras Literarias y Artísticas (modificada el 28 de septiembre de 1979), el Convenio de Roma de 1961, el Tratado de la OMPI de 1996 sobre Interpretación o Ejecución y Fonogramas, y la Convención Universal sobre Derecho de Autor (revisada el 24 de julio de 1971).