

Geolocalización de Nodos Inalámbricos

Matías Parodi, Guillermo Reisch

Abstract—Cada vez son más los dispositivos móviles con servicios y aplicaciones sensibles a la posición del usuario. Para esto en general se usa GPS consiguiendo excelentes resultados. En el interior de muchos edificios su utilización se ve fuertemente limitada debido a que en general la señal de los satélites que intervienen en la triangulación no es muy buena, pudiendo ser afectada por obstáculos e interferencias. En estos casos la precisión puede ser mejorada haciendo uso de sistemas de geolocalización que utilizan información de redes inalámbricas de corto alcance como WiFi, una técnica que es posible gracias a la proliferación de routers y access points cuya posición aproximada sí se conoce. El enfoque presentado parte de la base que se tiene la posición de al menos tres routers inalámbricos presentes en el entorno. Se estudiará como mediante un algoritmo evolutivo se puede aproximar la posición de un dispositivo móvil, si este es capaz de medir las intensidades de las señales hacia cada uno de dichos transmisores.

Index Terms—Algoritmos Evolutivos; triangulación; WiFi

I. INTRODUCCIÓN

EL presente informe consiste en la evaluación del comportamiento de un algoritmo evolutivo aplicado a la triangulación de dispositivos móviles en relación a la posición de al menos tres routers inalámbricos, cuyas posiciones absolutas o relativas a un centro arbitrario sí se conocen.

Se asumirá que los routers presentes en el entorno comparten características similares (e.g. alcance máximo), lo cual resulta razonable asumir para un gran número de routers WiFi existentes en el mercado. Esto será considerado al momento de la inicialización de la población inicial.

Por otro lado, resulta inapropiado asumir que las señales no se ven afectadas por obstáculos o fuentes de interferencia encontrados en el entorno, lo cual obligará a estimar, además de la posición, cuál es la caída o atenuación de la señal hacia cada router y para esto se estudiarán varias alternativas de modelo de propagación de ondas.

19 Julio, 2013

II. MOTIVACIÓN

Debido a la explosión de las redes WiFi en todo el mundo [1], en los últimos años han surgido servicios de geolocalización por triangulación WiFi [2] que ofrecen APIs [3] y SDKs [4] para el desarrollo de aplicaciones sensibles a la posición del usuario.

Desafortunadamente, ninguna de las grandes empresas en estas áreas han publicado especificaciones de sus implementaciones, por lo que se desconocen muchos detalles

del funcionamiento interno y en particular los algoritmos y técnicas que utilizan.

En el ámbito científico se exploraron varias alternativas. Quizás el enfoque más sencillo consiste en la triangulación de nodos a partir simplemente de las intensidades de las señales a cada router inalámbrico, como se analiza en [5]. Este enfoque presenta varias desventajas ya que no considera muchos aspectos de la realidad como la caída de señal por ruido electromagnético u obstáculos en el ambiente, condiciones climatológicas que afectan la señal, etc.

Otro grupo de investigadores han explorado el problema desde el área de machine learning, consiguiendo buenos resultados utilizando dos enfoques distintos. Por un lado, se han implementado satisfactoriamente sistemas basados en *fuzzy logic* [6], mientras que otros han intentado la integración de más de un sistema de triangulación, con el fin de tomar los beneficios de cada uno de ellos, combinando los resultados a partir de constantes de peso aprendidas, que se van ajustando a medida que se ejecuta el algoritmo [7].

Dado el tamaño del espacio de búsqueda, resulta interesante atacar el problema mediante un algoritmo evolutivo con el objetivo de lograr aproximar la posición con una precisión comparable a los servicios existentes pero para eso habrá que restringir el problema de lo contrario sería imposible abordarlo dada la cantidad de variables libres presentes.

III. DIFICULTADES

El nivel de señal recibido por el dispositivo que se intenta triangular varía con el tiempo, aunque no hayan cambiados sus posiciones relativas. Es importante entender qué ocasiona dichas fluctuaciones, ya que tener una buena medición de la señal será esencial para la correcta evaluación del individuo.

Por mayores detalles respecto a cuales son los factores más importantes que afectan a las señales inalámbricas recibidas y como estos influyen en la misma referirse a [6] y [8].

A. Tecnología

Las cuatro características más importantes que caracterizan a los dispositivos con los que se trabaja son: la frecuencia de transmisión, el tipo de antena, la ganancia de la antena y la potencia de transmisión.

A pesar de que existen muchas marcas y modelos existentes en el mercado, los fabricantes de routers modernos

respetan estándares internacionales para garantizar el correcto funcionamiento de redes WiFi con cualquier dispositivo inalámbrico.

Mediante un estudio de mercado se encontró que los routers domésticos son bastante homogéneos en cuanto a sus características, lo cual nos permite asumir algunas cosas que ayudarán a simplificar el problema de manera de hacerlo abordable.

La frecuencia de transmisión depende del canal en el que se encuentre transmitiendo el router, sin embargo el rango en el que puede estar transmitir no es demasiado grande, pudiendo variar entre 2,412GHz a 2,472GHz. Para simplificar el problema se asumirá entonces que el router opera en una frecuencia similar al promedio, es decir 2,442GHz, correspondiente al canal 7 [9].

En cuanto al tipo de antena, se encontró que todos los routers modernos tienen entre una y dos antenas isotrópicas de ganancia 5dBm, que emiten uniformemente hacia todos lados, por lo que la energía se disipará en el ambiente proporcionalmente al cuadrado de la distancia.

Por otro lado muchos routers permiten configurar la potencia de transmisión, pudiendo variar entre 1dBm (1mW) y 20dBm (100mW). A pesar de esto en la práctica la mayoría de los routers usan la potencia máxima de transmisión, la cual en general está activada por defecto, con el fin de hacer llegar la señal más lejos.

B. Entorno

El entorno puede afectar significativamente la calidad e intensidad de la señal, en particular el ruido electromagnético, condiciones climáticas y obstáculos presentes ocasionan una pérdida significativa de la misma.

Dado que es imposible construir un modelo de propagación para cada entorno, ya que este afecta de manera diferente a cada dispositivo, resulta vital estimar la caída de la señal percibida por el dispositivo móvil, en lugar de intentar modelar cada variable por separado.

Se considerará que la señal real es igual a la señal emitida por el router, menos la pérdida por disipación (dependiente de la distancia), menos la pérdida por obstáculos e interferencias (dependiente del entorno).

C. Muestreo

Una tarjeta WiFi estándar solo puede escuchar un canal por vez, por lo que para obtener todas las redes visibles en un entorno es necesario iterar en todos los canales en busca del anuncio de los routers inalámbricos.

Esto ocasiona que la información obtenida sobre una red en un determinado momento esté desactualizada por algunos segundos. Esto, sumado a los factores antes mencionados, ocasiona que mediciones consecutivas pueden dar valores de intensidad de señal bastante distintos.

Para mitigar este problema se puede considerar como intensidad de señal hacia un router dado el promedio de las últimas 5 lecturas.

D. Cantidad de Routers

La cantidad de routers afectará la precisión con la que se consigue triangular al dispositivo móvil pero en cualquier caso serán necesarios al menos tres, de lo contrario no será posible determinar su posición en el plano.

IV. MODELOS EXISTENTES

El modelo de propagación de ondas permitirá la conversión de una intensidad de señal recibida S_r (equivalentemente, una potencia de recepción P_r) a una distancia D y viceversa, a través de una función $signal(D)$ y su inversa $distance(S_r)$.

Encontrar una buena aproximación de $distance$ es esencial ya que a partir de esta será posible encontrar una aproximación del error entre la distancia real y la distancia estimada del dispositivo.

Existen muchas aproximaciones que fueron analizadas en [10], [8] y [11], aquí se describirán las cuatro más representativas con el objetivo de analizar distintos enfoques.

A. Empirical Propagation Model

Este modelo de propagación es quizás el más preciso de todos a pesar de su simplicidad. La principal desventaja, que lo hace impracticable en muchos casos, es que requiere del estudio preciso del ambiente en cuestión, pero como vimos antes esto depende fuertemente de las condiciones en un instante dado, por lo que su utilidad es limitada a medida que avanza el tiempo.

$$distance(S_r) = |S_r - S_t| \times K$$

En donde S_t es la intensidad de señal transmitida por el router y K es un valor empírico que corresponde a la distancia cubierta por 1dBm.

B. Wall Attenuation Factor Propagation Model

Este modelo es una adaptación del *Floor Attenuation Factor Propagation Model* sugerido en [12].

En lugar de considerar la atenuación de la señal debido al factor del piso del edificio, considera el efecto causado por

las paredes entre el transmisor y receptor.

La ecuación que relaciona la potencia en función de la distancia es:

$$P_r = P_0 - 10 \times n \times \log\left(\frac{D}{D_0}\right) - \begin{cases} n \times W \times WAF & nW < C \\ C \times WAF & nW \geq C. \end{cases}$$

En donde n indica el ratio de incremento de la atenuación debido a la distancia, P_0 es la potencia de la señal a una distancia de referencia D_0 , C es la cantidad de paredes hasta la que el factor de atenuación se ve afectado y WAF es el factor de atenuación.

Como se puede apreciar este modelo es bastante más complejo y sigue requiriendo un análisis empírico para determinar el factor de atenuación.

C. Land Propagation Model

Un tercer modelo teórico bastante más complejo considera la pérdida de potencia en la señal debido a tres factores principales:

- *Path loss* (L_p): se refiere a la pérdida promedio debido a la propagación de las ondas en el espacio
- *Slow fading* (L_s): se refiere a pérdidas de señal causadas en distancias largas debido a variaciones lentas en el espacio-tiempo [13]
- *Fast fading* (L_f), se refiere a pérdidas de señal por variaciones bruscas en el espacio [14] [15].

De esta forma la ecuación que relaciona potencia con distancia está dada por:

$$P_r = \frac{G_r \times G_t}{L_p \times L_s \times L_f} \times P_t$$

En donde G_r y G_t son las ganancias de las antenas de recepción y transmisión respectivamente y P_t es la potencia con la que transmite el router.

Calcular L_p , L_s y L_f es complejo y requiere de mediciones del ambiente en donde se va a desenvolver el algoritmo, lo que lo hace poco práctico en este contexto.

D. Free Space Propagation Model

Este modelo de propagación está inspirado en la ecuación de propagación de *Friis* [16] y es quizás el más útil de todos ya que ofrece un compromiso entre simplicidad y utilidad.

La ecuación en cuestión es:

$$P_r = \frac{G_t \times G_r \times \lambda^2}{(4\pi \times D)^2} \times P_t$$

En donde λ es la longitud de onda usada.

A partir de esta ecuación se puede despejar la distancia teórica entre ambos nodos, resultando:

$$d(L_{dB}) = \sqrt{\frac{G_t \times G_r \times \lambda^2}{(4\pi)^2 \times P_r}} \times P_t$$

λ se puede obtener a partir de la frecuencia F y la velocidad de la luz C :

$$\lambda = \frac{C}{F}$$

Por otro lado, usando las fórmulas de decibel se puede obtener la potencia P_r en función de la intensidad de la señal L_{dB} y una potencia de referencia P_0 , que en el caso de las redes inalámbricas es 1mW.

$$P_r = 10^{\frac{L_{dB}}{10}} \times P_0$$

A partir de esto es posible reescribir la ecuación que buscamos de la siguiente forma:

$$d(L_{dB}) = \sqrt{\frac{G_t \times G_r \times \left(\frac{C}{F}\right)^2}{(4\pi)^2 \times 10^{\frac{L_{dB}}{10}} \times P_0}} \times P_t$$

Simplificándola y usando los valores antes mencionados se obtiene finalmente una ecuación de la distancia entre el transmisor y receptor en función de la intensidad de señal recibida:

$$d(L_{dB}) = \sqrt{\frac{5 \times 5 \times \left(\frac{299.792.458}{2.442.000.000}\right)^2}{(4\pi)^2 \times 10^{\frac{L_{dB}}{10}} \times 0,001}} \times 0,1 = \sqrt{\frac{0.2386}{10^{\frac{L_{dB}}{10}}}}$$

V. ALGORITMO

A. Tecnologías Usadas

Para la implementación del algoritmo se decidió usar los lenguajes de programación C/C++ [17] [18] debido a su sencillez, bajo consumo de memoria, velocidad de ejecución y herramientas disponibles [19].

Se implementará el algoritmo usando el framework *Malva* [20], el cual es un fork de *Mallba* [21]. Se estudiaron otras alternativas pero finalmente se optó por este debido a recomendaciones hechas en el curso y al hecho de que contábamos con el soporte del foro de discusión ante cualquier inconveniente.

B. Variante

Luego de estudiar varias alternativas se optó por usar la variante de *malva* denominada *GA*, correspondiente a un algoritmo genético genérico con operadores personalizados, con el objetivo de tener total control sobre la aplicación de los mismos sobre los individuos.

C. Representación Interna

Para la representación interna del individuo se pueden adoptar dos enfoques distintos, dependiendo de como se interprete la intensidad de las señales leídas por la tarjetas de red inalámbrica del dispositivo.

En un comienzo se implementó el algoritmo utilizando el primero de dichos enfoques, pero luego de un estudio más profundo del modelo de propagación de ondas elegido, se encontró que el segundo enfoque ofrece ventajas sustanciales, por lo que se prosiguió a adaptar el algoritmo de acuerdo a él.

Para entender ambos enfoques es importante primero presentar dos escenarios de ejemplo con el objetivo de explicar de que manera afectan las dos alternativas al problema presentado.



Fig. 1: Escenario ideal

La imagen 1 presenta un escenario ideal en donde se tiene línea de visión entre el router transmisor R y el dispositivo receptor D , ubicados a una distancia de 10m uno del otro. La señal percibida en este ejemplo es de -25dBm.



Fig. 2: Escenario real

La imagen 2 presenta un escenario un poco más realista, en donde la línea de visión se ve obstruida por un obstáculo O (e.g. una pared) entre ellos. En este caso las ondas se ven perturbadas y el dispositivo solo percibe una señal de -30dBm, es decir, se tiene una atenuación de 5dBm.

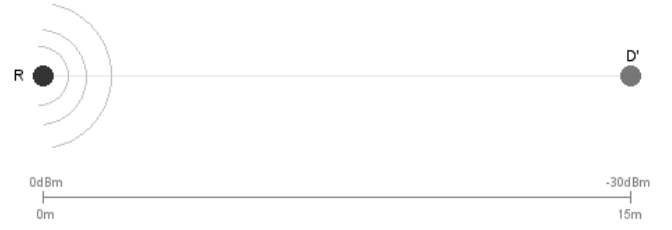


Fig. 3: Escenario percibido

La imagen 3 representa la posición D' en donde el dispositivo se percibiría a sí mismo de no tener en cuenta posibles atenuaciones. Una señal de -30dBm será interpretada por él como una distancia de 12m en lugar de 10m, pero esto como no es real. El problema de fondo es la imposibilidad de determinar la existencia de obstáculos entre ambos extremos y como este afecta a la señal.

1) *Enfoque 1 - Atenuación Explícita:* Cada individuo es una n -úpla de la forma:

$$I = (x, y, a_1, \dots, a_n)$$

$$x, y \in (-\infty, +\infty)$$

$$a_i \in [0, 1] \forall i \in (1, \dots, n)$$

En donde (x, y) se corresponde con las coordenadas estimadas del individuo en el plano y los valores a_i representan la atenuación estimada de la señal entre el dispositivo y el router i .

De acuerdo a la representación descrita $(10, -5, 2.7, 0.0, 1.1)$ representa un individuo en una posición estimada $(10, -5)$ en el plano (con respecto al origen de coordenadas arbitrario usado para inicializar los routers) que tiene un decaimiento estimado de señal de 2.7dBm, 0.0dBm y 1.1dBm con respecto al router R_1 , R_2 y R_3 respectivamente.

Si la intensidad de la señal que percibe de, por ejemplo, el router R_1 es de -30dBm, para el cálculo del fitness se asumiría que la misma debió ser 2.7dBm mayor, por lo cual se usará una señal teórica de $-30 + 2.7 = -27.3\text{dBm}$ en su lugar.

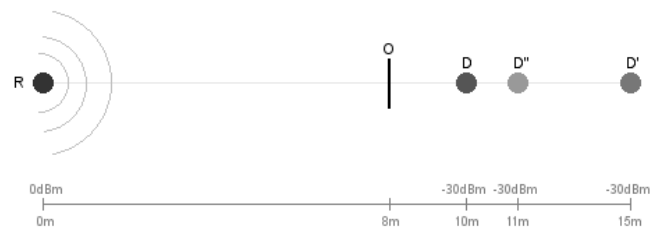


Fig. 4: Escenario mejorado

La imagen 4 muestra el efecto que tiene sobre la estimación

de la posición el hecho de considerar la atenuación hacia el router.

Al alterar la señal a partir de la atenuación estimada, el dispositivo D se acerca a la posición real D .

Este enfoque, si bien en un primero momento parecía adecuado, tiene un defecto importante. La representación interna del individuo contiene un valor de atenuación hacia cada uno de los routers, por lo que dado cualquier valor para (x, y) , el algoritmo será capaz de modificar dichos valores de manera que el fitness sea bueno. En otras palabras, sin importar la estimación de la posición, existirá un vector de atenuación que haga que el individuo sea aceptable, potencialmente mejor que otro cuya posición estimada es más acertada. Asimismo, la representación interna dependerá directamente de la cantidad de routers usados ya que para cada uno de ellos se deberá estimar un decaimiento, ocasionando que rápidamente el espacio de búsqueda crezca al punto de que no se explore lo suficiente en tiempos razonables.

2) *Enfoque 2 - Atenuación Implícita*: Cada individuo es un par de la forma:

$$I = (x, y) \\ x, y \in (-\infty, +\infty)$$

En donde (x, y) se corresponde con las coordenadas estimadas del individuo en el plano.

Este segundo enfoque extiende al anterior haciendo un mejor uso de la información obtenida del modelo de propagación de ondas. La idea a grandes rasgos consiste en inferir cuál es el decaimiento debido a los obstáculos, de manera que no sea necesario que el algoritmo lo estime.

Para esto es importante notar que dado cualquier punto en el plano, existen dos valores de intensidad de señal. Por un lado se tiene la *señal teórica*, que corresponde a la señal que debería estar llegando en un escenario ideal y queda determinada únicamente a partir de la distancia entre el transmisor y el receptor. Por otro lado se tiene la *señal recibida*, que corresponde a la *señal teórica* menos un *decaimiento* debido a obstáculos e interferencias.

De esta manera, la diferencia de ambas magnitudes es una buena aproximación del decaimiento usado en el enfoque anterior.

$$a_i = \text{signal}(\sqrt{(R_x - x)^2 + (R_y - y)^2}) - S_r$$

Este enfoque no requiere una representación interna dependiente de la cantidad de routers y solo se limita a ajustar la posición del dispositivo, logrando un espacio de búsqueda mucho más acotado, aunque infinito.

La idea de aquí en más será adoptar esta alternativa y minimizar el vector de decaimientos inferido, o

equivalentemente, minimizar la suma de la diferencia entre la distancia estimada (calculada a partir de la posición (x, y)) y la distancia teórica (calculada a partir de la señal percibida) respecto a cada router.

Es importante recordar que esto se estará haciendo para más de un router al mismo tiempo, por lo que el algoritmo deberá encontrar un punto de equilibrio. Si bien el vector del óptimo global será mínimo, en principio no será nulo.

D. Función de Fitness

La función de fitness a minimizar es la suma de errores respecto a la distancia estimada y la distancia teórica (calculada usando el modelo de propagación de ondas *free space propagation model*).

$$f(x, y) = \sum_{i=1}^n |\sqrt{(R_{i_x} - x)^2 + (R_{i_y} - y)^2} - \text{distance}(R_{i_s})|$$

En donde (R_{i_x}, R_{i_y}) y R_{i_s} son la posición y la intensidad de señal recibida del router i respectivamente, mientras que (x, y) es la posición estimada del dispositivo que se está triangulando.

La función de fitness así expresada es una aproximación de la suma de los errores reales. Si bien se espera que su valor se aproxime a 0, difícilmente llegará a anularse completamente ya que cualquier atenuación debido a un obstáculo o interferencia implica en algún sentido pérdida de información, imposibilitando lograr triangular al dispositivo con una precisión arbitrariamente grande.

E. Inicialización

Como se ha dicho antes, se asume que los routers comparten características similares, en particular, que el alcance máximo para todos ellos es de A metros (en la práctica se encontró que dicho valor es aproximadamente 50m [22]).

La población será inicializada eligiendo valores aleatorios para (x, y) de manera que todos los individuos comiencen dentro del cuadrado que circunscribe todos los círculos de centro (R_{i_x}, R_{i_y}) y radio A .

De esta manera, durante la inicialización de la población los valores estarán limitados por:

$$x_{min} = \min_{i=1} R_{i_x} - A \quad x_{max} = \max_{i=1} R_{i_x} + A \\ y_{min} = \min_{i=1} R_{i_y} - A \quad y_{max} = \max_{i=1} R_{i_y} + A$$

F. Operadores

Dado que el espacio de búsqueda es bastante grande, resulta esencial mantener la diversidad hasta haber muestreado una amplia área del plano, ya que en principio se puede asumir la existencia de muchos óptimos locales. Dicho esto, los operadores que serán utilizados se describen a continuación.

1) *Selección*: Se usará selección por torneo, habiendo probado también selección por ranking y roulette, obteniendo en todos los casos resultados similares.

2) *Cruzamiento*: El operador de cruzamiento promediará los valores de ambos padres.

$$x_3 = \frac{x_1 + x_2}{2} \quad y_3 = \frac{y_1 + y_2}{2}$$

3) *Mutación*: El operador de mutación alterará la posición de acuerdo a un valor aleatorio elegido en un determinado rango.

$$x = x + k_1 \quad y = y + k_2 \\ k_1, k_2 \in (n, m)$$

4) *Búsqueda Local*: Las distancias con las que se trabajan son del orden de unos pocos metros a algunas decenas de metros. Dadas estas magnitudes, un error de entre 5cm y 25cm son igual de aceptables. Por lo tanto no existe la necesidad de optimizar el individuo para alcanzar la mejor estimación posible en su entorno.

Se considerará que un individuo con errores de estos ordenes es suficientemente bueno, de modo que no será necesario intentar disminuir aún más dicho error.

G. Reemplazo de Individuos

Dadas las características del framework se decidió realizar el reemplazo poblacional luego de aplicado el operador de cruzamiento. Para esto se seleccionan los dos mejores individuos entre los padres y sus respectivos hijos, pudiendo ocurrir que el individuo descendiente reemplace a uno de sus padres si su fitness es mejor que el de ellos.

H. Criterio de Parada

A partir de lo dicho anteriormente parece razonable parar el algoritmo si se consigue un individuo cuya suma de errores (es decir, el valor de la función de fitness) es menor a 5cm.

Al mismo tiempo será necesario un límite en el número de generaciones con el objetivo de evitar que el algoritmo converja a un óptimo local en donde no puede lograr tal precisión, ocasionando que no termine de ejecutarse. Para evitar esto empíricamente se encontró que un límite de 100 generaciones es suficiente, ya que en la mayoría de las ejecuciones la generación del mejor individuo estaba por debajo de este límite.

I. Mecanismo de Reparación

Cabe señalar que los individuos son válidos desde el momento de su inicialización y los operadores aplicados no los corrompen, por lo que no hace falta un mecanismo de reparación o de descarte de individuos inválidos.

VI. ANÁLISIS EXPERIMENTAL

A. Instancias de Prueba

Las instancias de prueba son generadas de manera de simular entornos realistas.

Para esto se define la posición real del dispositivo y cierta cantidad de routers en un círculo de centro (0, 0) y radio 15m.

Una vez hecho se tiene la posición de todos los nodos presentes se calcula la distancia real entre el dispositivo y cada uno de los routers. Mediante esta distancia y el modelo de propagación de ondas se obtiene la señal real que percibiría, a la cual se le sustrae un valor aleatorio, correspondiente al decaimiento, de entre 0dBm y 2dBm.

B. Entorno de Prueba

Dado que en la práctica la implementación realizada no parece necesitar de grandes tiempos de ejecución ni grandes cantidades de memoria o procesamiento que requieran correrlo en un cluster, se usará una computadora de características moderadas para realizar las pruebas.

1) *Marca y Modelo*: Dell Inspiron N5010

2) *Procesador*: Intel Core i5 2.67GHz (doble nucleo, cuatro threads), 64KB de cache L1, 512KB de cache L2, 3MB de cache L3

3) *Memoria*: 4GB de memoria DDR3 1333MHz

4) *Sistema Operativo*: GNU/Linux [23] [24], en su variante Debian 7.0 Wheezy [25] para 64 bits (amd64).

Se prestó especial cuidado a los procesos ejecutándose, deteniendo todos las aplicaciones y servicios innecesarios al momento de realizar las pruebas.

C. Ajuste Paramétrico

Para el ajuste paramétrico se usará una instancia de prueba reducida, conformada por 3 routers ubicados aleatoriamente en el plano. A partir de dicha instancia y para cada combinación de parámetros se harán 3 ejecuciones independientes.

La tabla I muestra las combinaciones de parámetros estudiados para hacer el ajuste paramétrico, el fitness corresponde al mejor individuo encontrado en las 5 ejecuciones independientes.

TABLE I: Ajuste paramétrico

Población	Cruzamiento	Mutación	Fitness
5	0.70	0.01	0.4776
		0.05	0.2555
		0.10	0.3230
	0.80	0.01	0.3166
		0.05	0.3383
		0.10	0.2588
	0.90	0.01	0.2606
		0.05	0.2965
		0.10	0.2566
15	0.70	0.01	0.2707
		0.05	0.3181
		0.10	0.2558
	0.80	0.01	0.2623
		0.05	0.2568
		0.10	0.2561
	0.90	0.01	0.2701
		0.05	0.2605
		0.10	0.2553
30	0.70	0.01	0.2559
		0.05	0.2554
		0.10	0.2553
	0.80	0.01	0.3199
		0.05	0.2554
		0.10	0.2554
	0.90	0.01	0.3194
		0.05	0.2561
		0.10	0.2554

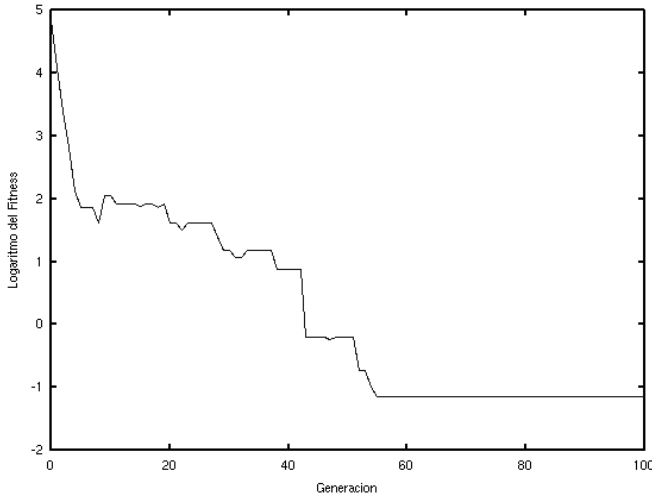


Fig. 5: Evolución del fitness en función de la generación

Análisis: A partir de los resultados obtenidos se puede apreciar como, para todas las combinaciones de parámetros con los que se probó, las soluciones son muy buenas. Existe cierta tendencia a obtener soluciones marginalmente peores para una probabilidad de cruzamiento 0.01, en especial combinando esto con poblaciones chicas, una posible explicación de este fenómeno es la diversidad.

Con el objetivo de obtener soluciones de calidad y al mismo tiempo mantener la diversidad en la población se optará de aquí en más por una población de 30 individuos, una probabilidad de 0.9 para el operador de cruzamiento y de 0.1 para el operador de mutación.

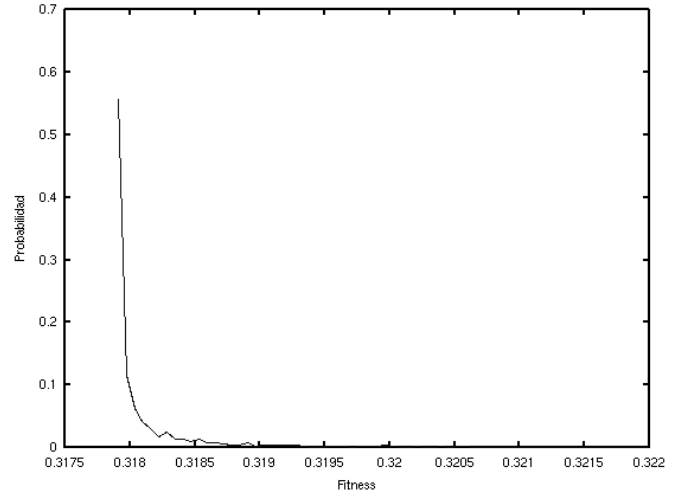


Fig. 6: Distribución de la probabilidad de fitness

Análisis de la Distribución: Se realizaron pruebas de normalidad y exponencialidad con el test de Kolmogorov–Smirnov, en ambos casos se obtuvo resultados negativos.

D. Evaluación

Para realizar la evaluación del algoritmo se usarán 25 instancias de prueba divididas en 5 conjuntos de 3, 4, 5, 6 y 7 routers. Para cada instancia se harán 5 ejecuciones independientes, completando en total 125 ejecuciones.

Para todos los casos se reportará:

- G : generación del mejor individuo (I)
- f : fitness de I
- f_{AVG} : promedio de fitness de los individuos de G
- σ_f : desviación estándar de los individuos de G
- t : segundos utilizados para encontrar I
- t_{AVG} : promedio de t para todas las ejecuciones
- σ_t : desviación estándar de t para todas las ejecuciones
- $Error$: error en metros respecto de la posición real

Análisis: Todas las soluciones obtenidas son muy buenas, presentando en general errores del orden de 1m. Por otro lado los tiempos de ejecución son muy cortos, evidenciando la capacidad del algoritmo implementado para hallar rápidamente una excelente aproximación de la posición real del dispositivo.

A partir del fitness de los mejores individuos y de las generaciones en las que fueron encontrados parece correcto afirmar que aunque se corriera el algoritmo por más tiempo no sería posible mejorar el error. De hecho, parece estar encontrando las mejores soluciones posibles, pero está sujeto a la falta de información debido a las atenuaciones existentes.

A medida que aumenta la cantidad de routers se mejora marginalmente los resultados, aún así cabe señalar la variación de errores para instancias de prueba con la misma cantidad de routers, lo cual es razonable ya que una correcta triangulación

TABLE II: Resultados

R	G	f	f_{AVG}	σ_f	t	t_{AVG}	σ_t	Error
3	63	0.801	0.814	0.043	0.003	0.003	0.001	2.324
	32	0.067	0.067	0.000	0.001	0.003	0.001	0.403
	37	0.124	0.134	0.059	0.002	0.003	0.001	0.470
	67	0.249	0.268	0.035	0.004	0.003	0.001	1.870
	93	0.931	0.940	0.026	0.004	0.003	0.001	0.276
4	68	0.544	0.564	0.093	0.003	0.004	0.001	3.711
	64	4.613	4.615	0.004	0.003	0.004	0.001	0.914
	66	0.903	0.991	0.326	0.003	0.004	0.001	1.761
	72	0.469	0.505	0.117	0.004	0.004	0.001	1.083
	73	0.538	0.538	0.000	0.004	0.003	0.001	1.282
5	84	1.515	1.627	0.375	0.005	0.005	0.001	1.961
	75	3.272	3.505	0.360	0.005	0.005	0.002	1.595
	97	1.919	2.281	1.051	0.006	0.004	0.001	1.198
	91	2.056	2.066	0.037	0.005	0.005	0.001	1.775
	78	1.318	1.366	0.171	0.005	0.004	0.001	2.932
6	91	1.964	2.192	0.513	0.006	0.005	0.002	0.224
	80	2.065	2.094	0.072	0.006	0.005	0.001	1.446
	56	3.902	3.976	0.402	0.004	0.006	0.001	1.668
	74	4.044	4.179	0.396	0.005	0.005	0.002	1.379
	63	0.389	0.429	0.220	0.005	0.005	0.002	2.472
7	46	1.637	1.645	0.014	0.004	0.006	0.002	0.957
	66	1.025	1.083	0.150	0.005	0.006	0.002	0.813
	63	1.897	1.898	0.000	0.005	0.006	0.002	1.209
	95	3.307	3.307	0.003	0.008	0.006	0.002	1.514
	58	5.460	5.484	0.083	0.005	0.006	0.002	1.378

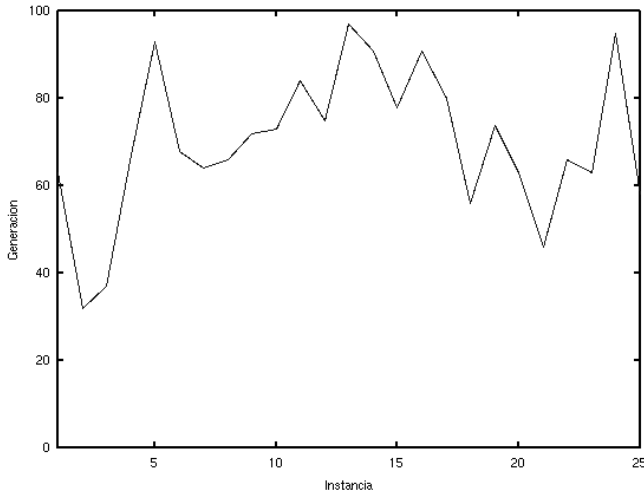


Fig. 7: Generación de los mejores individuos

depende en gran medida de la posición relativa de los nodos.

En general, parece primar la imprecisión intrínseca del modelo elegido, no pudiendo mejorar arbitrariamente la precisión, como se había predecido en un principio.

VII. CONCLUSIÓN

A juzgar por los resultados obtenidos en el análisis experimental se puede concluir la validez de la propuesta de usar algoritmos evolutivos para triangular nodos inalámbricos.

Como se evidencia para cada uno de los casos de prueba usados, el algoritmo fue capaz de encontrar una aproximación bastante precisa de la posición real del dispositivo. Si bien esta aproximación está limitada por incertidumbres en las

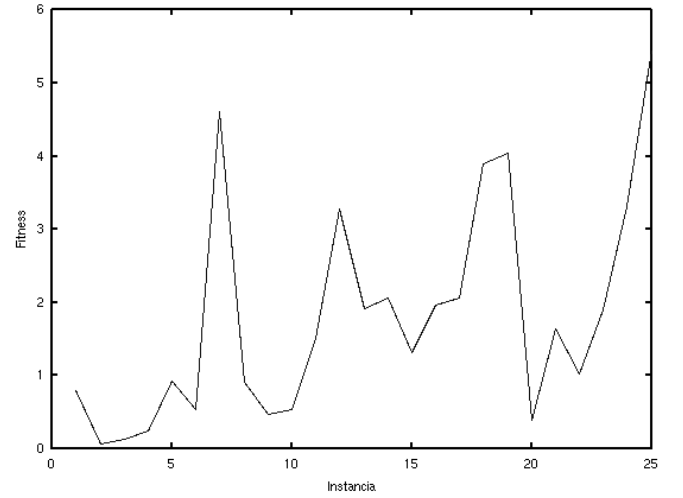


Fig. 8: Fitness de los mejores individuos

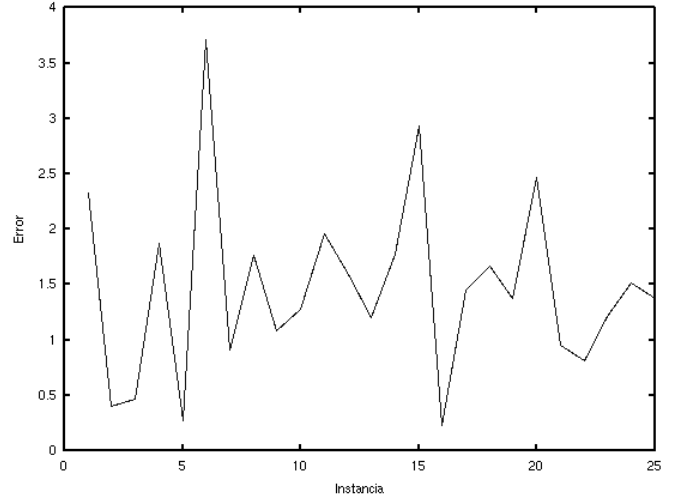


Fig. 9: Errores reales de los mejores individuos

mediciones y sesgada por el modelo de propagación de ondas usado, se demostró la posibilidad de triangular dispositivos móviles en tiempos muy reducidos.

Concluimos que un enfoque como el presentado aquí perfectamente puede ser usado en la práctica si se cuenta con al menos tres routers de los cuales sí se conoce su ubicación.

VIII. TRABAJO FUTURO

Durante el análisis del problema se hicieron varias suposiciones, que si bien son adecuadas en el contexto en el que se trabaja, no dejan de ser una simplificación de la realidad.

En el futuro se debería adoptar otro modelo de propagación de ondas más sofisticado con el objetivo de ver de que manera esto influye en las soluciones encontradas por el algoritmo.

Además de esto sería interesante triangular a los dispositivos en un espacio tridimensional, lo que requeriría agregar una nueva coordenada a la representación interna de los individuos y tener ciertas consideraciones especiales respecto a las antenas.

Por último, por no contar con el equipo adecuado y los medios para realizar mediciones suficientemente precisas se optó por limitarse a la simulación de triangulaciones con ejemplos generados aleatoriamente. En un futuro se deberá poner a prueba la propuesta utilizando routers reales en un ambiente controlado, con el objetivo de poner a prueba la validez del estudio realizado.

REFERENCES

- [1] Skyhook's mapped wifi access points. <http://www.skyhookwireless.com/location-technology/coverage.php>.
- [2] Wifi positioning system. http://en.wikipedia.org/wiki/Wi-Fi_positioning_system.
- [3] Google maps geolocation api. <https://developers.google.com/maps/documentation/business/geolocation>.
- [4] Skyhook location sdk. <http://www.skyhookwireless.com/location-technology>.
- [5] M. A. Ansari, S. Farooqui, and A. Yasar. Low cost solution for location determination of mobile nodes in a wireless local area network. pages 1–7, 2005.
- [6] T. Garcia-Valverde, A. Garcia-Sola, H. Hagra, J. Dooley, V. Callaghan, and J. Botia. A fuzzy logic based system for indoor localisation using wifi in ambient intelligent environments. *Fuzzy Systems, IEEE Transactions on*, PP(99):1–1, 2012.
- [7] Sheng-Cheng Yeh, Wu-Hsiao Hsu, and Yih-Shyh Chiou. Adaptive-weighting schemes for location-based services over heterogeneous wireless networks. pages 1–4, 2010.
- [8] P. Bahl and V.N. Padmanabhan. Radar: an in-building rf-based user location and tracking system. 2:775–784 vol.2, 2000.
- [9] A.B. Fadha, A. Al-Lawati, S. Al-Maskari, A. Touzene, and A. Al-Kindi. Experimental performance evaluation of wireless 802.11b networks. pages 151–155, 2008.
- [10] M. A. Ansari, S. Farooqui, and A. Yasar. Low cost solution for location determination of mobile nodes in a wireless local area network. pages 1–7, 2005.
- [11] R. Ezzine, A. Al-Fuqaha, R. Braham, and A. Belghith. A new generic model for signal propagation in wi-fi and wimax environments. pages 1–5, 2008.
- [12] S.Y. Seidel and T.S. Rappaport. 914 mhz path loss prediction models for indoor wireless communications in multifloored buildings. *Antennas and Propagation, IEEE Transactions on*, 40(2):207–217, 1992.
- [13] D. P. Agrawal and Q-A. Zeng. *Introduction to Wireless and Mobile Systems*. Thomson, 2006.
- [14] T.K. Sarkar, Zhong Ji, Kyungjung Kim, A. Medouri, and M. Salazar-Palma. A survey of various propagation models for mobile communication. *Antennas and Propagation Magazine, IEEE*, 45(3):51–82, 2003.
- [15] Theodore S. Rappaport. *Wireless Communications: Principles and Practice*. Prentice Hall, 2002.
- [16] Harald T. Friis. Introduction to radio and radio antennas. *Spectrum, IEEE*, 8(4):55–61, 1971.
- [17] The c resources network. <http://www.cprogramming.com/>.
- [18] The c++ resources network. <http://www.cplusplus.com/>.
- [19] D.A. Cargill and M. Radaideh. A practitioner report on the evaluation of the performance of the c, c++ and java compilers on the os/390 platform. pages 40–45, 2000.
- [20] Malva. <https://code.google.com/p/malva/>.
- [21] Gabriel Luque. Mallba library v2.0. <http://neo.lcc.uma.es/mallba/easy-mallba/>, 2006.
- [22] What is the range of a typical wi-fi access point? <http://help.unc.edu/help/what-is-the-range-of-a-typical-wi-fi-access-point/>.
- [23] The gnu operating system. <http://www.gnu.org/>.
- [24] The linux kernel archives. <https://www.kernel.org/>.
- [25] Debian wheezy release information. <http://www.debian.org/releases/stable/>.