

Análisis multiobjetivo de la selección de líderes en redes inalámbricas de sensores

Karen Miranda, Antonio López-Jaimes, Abel García-Nájera

Universidad Autónoma Metropolitana (UAM),
Depto. de Matemáticas Aplicadas y Sistemas,
Cuaajimalpa, Ciudad de México, México

{kmiranda,alopez,agarcian}@correo.cua.uam.mx

Resumen. Una red sensora consiste en un conjunto de sensores para monitorizar las condiciones ambientales. Los sensores transmiten cooperativamente la información recolectada a través de la red a una estación base que se encarga de recolectar la información de todos los sensores. Para ahorrar energía, los sensores se agrupan y eligen a un líder para que se encargue de reenviar los datos recolectados a la estación base. La selección del líder puede hacerse de manera aleatoria o con criterios específicos como energía residual, la distancia o la conectividad; así, el problema de agrupamiento y selección de líder es un problema NP-difícil. En este trabajo se analiza el problema de la selección de líderes de grupo desde una perspectiva multiobjetivo para determinar la pertinencia de algunos los objetivos utilizados que se pueden encontrar en la literatura.

Palabras clave: Optimización multi-objetivo, algoritmos genéticos, redes inalámbricas de sensores, cluster head.

A Multi-objective Analysis of the Cluster-head Selection Problem in WSN

Abstract. A Wireless Sensor Network (WSN) is composed of a set of energy and processing-constrained devices that gather data about a set of phenomena. An efficient way to enlarge the lifetime of a wireless sensor network is clustering organization, which structures hierarchically the sensors in groups and assigns one of them as a cluster-head. Such a head is in charge of specific tasks as data gathering from other cluster sensors and resending that data through the network to the base station. The cluster-head selection may be random or based on well defined criteria such as residual energy, node distance, signal strength or connectivity; hence, this problem is NP-hard for WSN. In the literature we may find proposals based on heuristics algorithms that consider at the same time different objectives and purposes. Therefore, in this paper we perform an analysis from the many-objective perspective about the pertinence of the most commonly used objectives.

Keywords: Multi-objective optimization, genetic algorithms, wireless sensor networks, cluster head.

1. Introducción

Las redes inalámbricas de sensores (Wireless Sensor Networks) están compuestas por un gran número de pequeños dispositivos electrónicos que son capaces de monitorizar condiciones ambientales, tales como, la temperatura, la humedad o la velocidad del viento. Los sensores recolectan datos sobre un fenómeno dado y reenvían esos datos a una estación base a través de la red que los mismos sensores forman. Estos dispositivos son de bajo costo y fáciles de desplegar; sin embargo, también tienen capacidad de memoria y de energía limitada. Así, se busca que la red sensora esté activa y trabajando el mayor tiempo posible por lo que es necesario que cada sensor administre sus recursos de forma autónoma y que permita el ahorro de energía [6].

Una manera de optimizar el tiempo de vida una red sensora es organizar jerárquicamente a los sensores, formando grupos o *clusters* y seleccionando un líder por cada cluster (o *cluster head* del inglés) [14]. Los líderes están a cargo de concentrar la información que los otros sensores han recolectado y a su vez la reenvía a la estación base. El agrupamiento es una técnica eficiente para ahorrar energía; paradójicamente, al llevar a cabo estas tareas adicionales, los líderes gastan más rápidamente su propia energía, por lo que cada vez que un líder se queda sin batería, es necesario elegir un nuevo líder de entre los sensores restantes.

La selección de los líderes puede seguir diferentes criterios como la energía residual, la distancia entre sensores, la potencia de la señal o la conectividad e incluso una combinación de criterios. Estos criterios se utilizan en función de los objetivos que se desean optimizar, por ejemplo, maximizar el tiempo de vida de la red o minimizar el consumo de energía. Este problema de selección se puede ver como un proceso de toma de decisiones multicriterio. Dependiendo del propósito particular de la red sensora, i.e., en la agricultura, monitorización de la vida silvestre o condiciones climatológicas podemos encontrar una gran variedad de objetivos que se desean optimizar [9]. Por lo tanto, es importante determinar cuáles de esos objetivos son relevantes para el problema, cuáles son deseables y cuáles se satisfacen al optimizar otros objetivos ya que se sabe que no todos los objetivos son necesarios [4]. En general, el objetivo primordial es maximizar el tiempo de vida de la red.

En este trabajo se presenta un análisis de tres de los objetivos que se encuentran comúnmente en la literatura [2]: *i)* minimizar la distancia entre miembros del grupo; *ii)* minimizar la distancia de los líderes a la base y; *iii)* maximizar la energía residual de los líderes. Para ello, realizamos un estudio experimental utilizando el bien conocido algoritmo NSGA-II y donde analizamos la formación de grupos de sensores y la relación de conflicto entre los objetivos.

El resto del documento está estructurado de la siguiente manera. En la Sección 2 presentamos el problema de selección de líder en el contexto de las

redes inalámbricas de sensores así como el modelo del sistema y de energía. En la Sección 3, describimos las bases de problemas multiobjetivo, conflicto entre objetivos y presentamos el problema de selección de líder como un problema multiobjetivo. En la Sección 4, recordamos los conceptos básicos del NSGA-II, así como la representación de las soluciones como cadenas binarias. Presentamos los resultados del estudio experimental en la Sección 5 y finalmente, concluimos este trabajo en la Sección 6.

2. Problema de selección de líder

Las redes inalámbricas de sensores son ampliamente usadas en la actualidad. Entre algunas de las aplicaciones se encuentran la agricultura de alta precisión, la motorización de volcanes, mapeo de la biodiversidad, la domótica (automatización de casas y edificios) y en la concepción de ciudades inteligentes [11]. Las redes de sensores se pueden clasificar por su propósito y sus capacidades físicas específicas. En particular, en este trabajo nos enfocamos a las redes donde los sensores se despliegan una vez y permanecen en el mismo lugar hasta que termina su ciclo de vida.

Dado que el objetivo de una red sensora es proveer información sobre diferentes fenómenos y muchas veces se despliegan en zonas de difícil acceso, se busca que dicha red cumpla su objetivo durante el mayor tiempo posible. Existen diferentes técnicas de ahorro energético para redes con estas características, por ejemplo, reducción de paquetes de datos o control de topología [3]. Particularmente, el agrupamiento (del inglés *clustering*) es un método para reducir la cantidad de datos transmitidos a través de la red hacia una estación base donde se caracterizan los datos y se obtiene información.

Este método consiste en formar grupos de sensores y entre ellos elegir uno que funja como líder de cada grupo [14]. El líder es el encargado recolectar los datos en el grupo y determina cuales de ellos son relevantes para ser reenviados a la estación base, reduciendo así la cantidad de datos intercambiados entre sensores y por consecuencia ahorrando energía [12].

Existen diferentes propuestas para el agrupamiento y la selección de líderes. Por ejemplo, los más sencillos únicamente consideran la energía residual de los nodos y eligen como líder a aquel con la mayor energía. Si bien es una regla simple que no consume recursos suplementarios para su ejecución, no siempre es útil para las aplicaciones específicas de las redes sensoras. De hecho, los objetivos a considerar para seleccionar a los líderes depende, precisamente, de la propia aplicación. Algunos ejemplos de objetivos, no restrictivos, son: escalabilidad, tolerancia a fallas, balance de carga, tiempo de vida de la red o incrementar la zona de cobertura de los sensores [2].

En general, se puede considerar un problema genérico de optimización multiobjetivo para asignación de recursos con entradas, salidas requeridas, objetivos y restricciones [9]. Así, se pueden ver reflejados objetivos específicos como maximizar el tiempo de vida de la red, maximizar la cobertura, minimizar el costo, minimizar el consumo de energía o maximizar la utilización del espectro.

2.1. Modelo de la red de sensores

Una red de sensores se representa como un grafo $G = V, E$, donde V es un conjunto finito de n sensores V_i más una estación base V_B y E es el conjunto finito de m conexiones entre sensores en la red. Por conveniencia, se asume que: *i)* los sensores son distribuidos aleatoriamente sobre un espacio cuadrado de dos dimensiones; *ii)* los sensores son estáticos una vez que son desplegados; *iii)* la energía de los sensores no es renovable; *iv)* inicialmente todos los sensores tienen la misma cantidad de energía; *v)* los enlaces de comunicación son bidireccionales; *vi)* las capacidades computacionales son las mismas para todos los sensores; *vii)* los sensores no conocen su ubicación precisa.

2.2. Modelo de energía

Dado que la mayoría de los objetivos en el análisis del consumo de energía de los líderes con base en la distancia entre los miembros del grupo y la distancia entre el líder y la estación base, es necesario establecer el modelo de energía utilizado en los casos de estudio enfocado a los métodos de agrupamiento. Para ello, usamos la notación y valores de los parámetros usados para nuestras evaluaciones como se describe en la Tabla 1 y el modelo de energía como sigue [13]:

Tabla 1. Notación.

Significado	Símbolo	Valor
Energía Inicial	E_I	9 J
Nodos	N	20–500
Número de Miembros en el <i>cluster</i>	N_M	
Miembro	M	100 μ J
Líder	LG	
Estación Base	EB	Posición (0,0)
Tamaño del mensaje	k	2000 bits
Energía usada para transmitir	E_{tx}	
Consumo electrónico	E_{elec}	50 nJ/bit
Energía usada por líder de grupo por miembro por mensaje	E_{LGM}	
Distancia al líder	D_{LG}	
Distancia a la EB	D_{EB}	
Energía por transmisión múltiple	ϵ_{amp}	
Área	(m \times m)	400 \times 400

3. Problemas de optimización multiobjetivo

Muchos problemas reales tienen no sólo una función objetivo, sino m de ellas. En este caso, tenemos el vector $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))$ de m funciones objetivo, en donde $\mathbf{f} : \mathcal{X} \rightarrow \mathbb{R}^m$ define el espacio objetivo. Entonces, tenemos que

minimizar: $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))$ sujeto a $g_i(\mathbf{x}) = 0$, $i = 1, \dots, p$, $h_j(\mathbf{x}) \leq 0$, $j = 1, \dots, q$.

Los principales problemas de optimización pertenecen a la clase de complejidad NP-difícil, esto quiere decir que no se conoce un método exacto que pueda resolver todos los casos del problema de forma eficiente. No obstante, los métodos heurísticos, a pesar de no ofrecer una garantía de desempeño, usualmente encuentran soluciones con ciertos criterios de calidad en un tiempo razonable. Se sabe que para los problemas que involucra agrupar miembros, específicamente el agrupamiento de sensores, son problemas de clase NP-difícil [8].

3.1. Conflicto entre objetivos

Una propiedad importante de un problema multiobjetivo es el conflicto entre sus objetivos. Si los objetivos no tienen conflicto entre ellos, entonces podríamos resolver el problema optimizando cada objetivo de manera independiente. Sin embargo, en algunos problemas, aunque existe cierto conflicto, entre algunos objetivos no hay conflicto. Aunque diferentes autores han propuesto una definición para conflicto (véase por ejemplo, [4, 5]), en este documento utilizamos la definición propuesta por Carlsson y Fullér [5], que es la más intuitiva.

Definición 1. Sea $S_{\mathcal{X}}$ un subconjunto de \mathcal{X} , entonces, dos objetivos pueden estar relacionados de las siguientes maneras (suponiendo minimización):

1. f_i está en conflicto con f_j en $S_{\mathcal{X}}$ si $f_i(\mathbf{x}^1) \leq f_i(\mathbf{x}^2)$ implica que $f_j(\mathbf{x}^1) \geq f_j(\mathbf{x}^2)$ para toda $\mathbf{x}^1, \mathbf{x}^2 \in S_{\mathcal{X}}$.
2. f_i apoya a f_j en $S_{\mathcal{X}}$ si $f_i(\mathbf{x}^1) \geq f_i(\mathbf{x}^2)$ implica $f_j(\mathbf{x}^1) \geq f_j(\mathbf{x}^2)$ para toda $\mathbf{x}^1, \mathbf{x}^2 \in S_{\mathcal{X}}$.
3. f_i y f_j son independientes en $S_{\mathcal{X}}$, en otro caso.

En los casos 2 y 3, estos objetivos son llamados no conflictivos o redundantes.

3.2. Selección del líder como problema multiobjetivo

En una red de sensores, el objetivo que se busca mejorar en la mayoría de los casos es el tiempo de vida de la red. Sin embargo, este objetivo no se puede evaluar directamente ya que solamente hasta que se pone a funcionar la red sabemos el tiempo de vida real de la red.

Por esta razón, en varias propuestas [1, 10] se han utilizado otros objetivos que al parecer contribuyen a maximizar el tiempo de vida de la red. Entre estos objetivos podemos encontrar:

1. Minimizar la distancia de los líderes a la base, ya que estos sensores son los que envían más mensajes.
2. Minimizar la distancia de cada sensor a su líder, ya que evita que líderes potenciales consuman rápidamente su energía.
3. Maximizar la energía residual de los líderes, ya que de esta manera será menos frecuente tener que reemplazar a un líder.

4. Optimizar el balance de carga, lo cual ayuda a que los líderes necesiten ser reemplazados en tiempos similares.

En varios trabajos se utilizan estos objetivos de manera individual [9], mientras que en otros se utilizan simultáneamente [2]. Sin embargo, aún no se conoce si todos estos objetivos son necesarios para encontrar la solución óptima. Como se mencionó en la sección anterior, en ciertos problemas algunos objetivos no son necesarios o esenciales [4, 5].

Por este motivo, tomaremos los tres objetivos más utilizados para mejorar el tiempo de vida de una red de sensores con el fin de analizar si todos los objetivos son esenciales o hay alguno redundante en un planteamiento multiobjetivo. Los tres objetivos que adoptamos en el estudio son los tres primeros de la lista anterior.

4. Algoritmo para resolver el problema de selección

4.1. Algoritmo de optimización

Puesto que este estudio está enfocado al análisis de las características del problema, como algoritmo de optimización utilizamos uno de los algoritmos evolutivos más usados y conocidos, es decir, el *Non-dominated Sorting Genetic Algorithm II* (NSGA-II) [7]. Si bien este algoritmo ya no se considera del estado del arte, utilizando un número adecuado de individuos y generaciones podemos obtener soluciones representativas del frente de Pareto óptimo.

Este algoritmo utiliza una clasificación por capas de los individuos de acuerdo a su no dominancia. Para esto se obtienen los individuos que dominan al resto de la población. Este grupo de individuos conforma el primer frente no dominado y se les asigna un valor de aptitud de uno. Posteriormente, estos individuos son ignorados momentáneamente y se genera el segundo frente no dominado de individuos, a los cuales se les asigna una aptitud aumentada en uno. Este proceso continúa hasta que todos los individuos están clasificados.

4.2. Representación de las soluciones

Un parte importante para aplicar NSGA-II al problema de selección del líder es la representación de las soluciones y las suposiciones acerca del problema.

Para representar una solución utilizamos una cadena binaria $\mathbf{x} \in \{0, 1\}^N$, donde N es el número de sensores. De esta manera, tenemos que $x_i = 1$ si el i -ésimo sensor es líder, y $x_i = 0$ si el sensor es parte de un grupo (sensor miembro).

Para facilitar la notación consideremos que los índices de los líderes de grupo conforman el conjunto $L = \{i \mid x_i = 1\}$ y que la posición del i -ésimo sensor se denota mediante $\mathbf{s}^{(i)} \in \mathbb{R}^2$. Por otra parte, la posición de la base está dada por $\mathbf{s}^{(\text{base})}$.

Así, para determinar a qué grupo pertenece un sensor miembro, buscamos el líder más cercano. Es decir, el grupo $G(\mathbf{s}^{(i)})$ al que pertenece el sensor $\mathbf{s}^{(i)}$ ($i = 1, \dots, N$) está dado por

$$G(\mathbf{s}^{(i)}) = \arg \min_{j \in L} \|\mathbf{s}^{(i)} - \mathbf{s}^{(j)}\|,$$

donde $\|\cdot\|$ es la distancia Euclideana entre dos puntos. El conjunto de índices de los sensores que son miembros del grupo de un líder j se denota mediante $M_j = \{i \mid G(\mathbf{s}^{(i)}) = j, i = 1, \dots, N\}$. Nótese que con esta definición cada grupo tendrá al menos como único miembro a un líder. Así, nuestros objetivos se definen de la siguiente manera:

1. Minimizar distancia de los líderes L a la base:

$$\text{Min } f_1(\mathbf{x}) = \frac{1}{|L|} \sum_{j \in L} \|\mathbf{s}^{(j)} - \mathbf{s}^{(\text{base})}\|.$$

2. Minimizar distancia de los sensores a su líder de grupo:

$$\text{Min } f_2(\mathbf{x}) = \frac{1}{|L|} \sum_{j \in L} \left(\frac{1}{|M_j|} \sum_{i \in M_j} \|\mathbf{s}^{(i)} - \mathbf{s}^{(j)}\| \right).$$

3. Maximizar energía residual de los líderes de cada grupo:

$$f_3(\mathbf{x}) = \frac{1}{|L|} \sum_{j \in L} E_I - LG_j,$$

donde E_I y LG son los valores definidos en la Tabla 1.

Con esta representación utilizamos la cruce binaria de dos puntos y la mutación uniforme. Solamente tuvimos que hacer una ligera modificación para evitar soluciones con cero líderes. Al final de la cruce y mutación, si el resultado es la cadena binaria con ceros solamente, entonces en una posición aleatoria de la cadena se escribe un uno.

5. Estudio experimental

Para realizar los experimentos consideramos instancias del problema donde los sensores están ubicados de manera aleatoria en una cuadrícula usando los parámetros descritos en la Tabla 1. Para todas las instancias, excepto para 20 sensores, se realizaron simulaciones usando el NSGA-II con 2000 individuos evolucionados durante 20,000 generaciones. El porcentaje de cruce utilizado fue 0.9, mientras que el porcentaje de mutación fue de $1/N$, donde N es el número de sensores. En el caso de 20 sensores fue posible obtener la solución óptima utilizando búsqueda exhaustiva.

Es importante mencionar que para facilitar el planteamiento del problema en NSGA-II, el objetivo para maximizar f_3 (la energía residual) fue transformado

a un objetivo de minimización. Es decir, en la discusión de las secciones siguientes el objetivo f_3 es minimizar la energía utilizada por los líderes de los grupos.

Para mostrar la forma y tipo de frente de Pareto que se produce al minimizar los tres objetivos se presenta la Fig. 1. La figura muestra el frente de Pareto obtenido por NSGA-II para una instancia con 300 sensores. Note que el tercer objetivo (eje z) se refiere a la energía utilizada por los líderes al recibir y enviar mensajes (la cual fácilmente se puede interpretar como energía residual). A diferencia de los frentes de Pareto típicos de problemas con 3 objetivos, en este caso el frente no es una superficie sino que parece ser una línea (objeto en una dimensión).

En las siguientes secciones se analizará con detalle el conflicto entre los objetivos del problema.

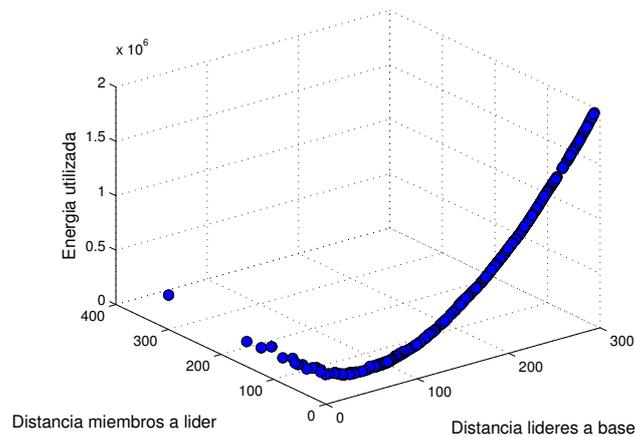


Fig. 1. Vista en 3D del frente de Pareto obtenido para una instancia con 300 sensores.

5.1. Formación de grupos de sensores

En esta sección se muestra la forma de los grupos de sensores y sus líderes según el compromiso entre los 3 objetivos del problema. En la Fig. 2 se ilustra el frente de Pareto óptimo compuesto por 82 soluciones para una instancia con solamente 20 sensores. Dado que para 20 sensores solamente hay $2^{20} \approx 10^6$ soluciones posibles, obtuvimos el conjunto de óptimos completo realizando una búsqueda exhaustiva al evaluar todas las soluciones posibles.

Para tener una idea del tipo de configuraciones de grupos de sensores y sus líderes, en las Fig. 3-5 se muestran 3 configuraciones de agrupaciones posibles que representan soluciones óptimas de Pareto de la Fig. 2. En primer lugar tenemos la solución que obtiene el mejor valor para la distancia del líder a la base (véase

Fig. 3). En esta figura se muestra la posición de los 20 veinte sensores utilizando círculos, mientras que el sensor líder es el círculo con un punto en su interior.

Recordemos que la base en todos los casos está localizada en la posición $(0, 0)$. Puesto que para esta pequeña instancia no hay conflicto entre este objetivo y la energía utilizada, la solución que minimiza la energía es la misma configuración de grupos. Es decir, la solución la solución óptima en términos de distancia a la base y energía es aquella que forma un único grupo donde el líder es el sensor más cercano al origen (la base).

En el otro extremo, tenemos a la solución que minimiza la distancia de los miembros del grupo a su líder. La Fig. 4 muestra los grupos conformados para esta solución óptima de Pareto. Como era de esperarse, para minimizar la distancia a su líder se forman grupos con un solo sensor, el cual es su propio líder. Como se muestra en la figura, hay 20 grupos denotados con un símbolo diferente y su líder es indicado mediante un punto en el centro. En otras palabras, esta solución es aquella donde no se utiliza el esquema de líderes ya que cada sensor transmite directamente a la base. Si bien esta configuración minimiza la distancia a los líderes también es la que más energía consume ya que la base está muy alejada para la mayoría de los líderes.

En la práctica, una solución que se podría elegir para implementar es alguna ubicada en la zona media del frente Pareto (llamada “rodilla” del frente en optimización multiobjetivo). La distribución de los grupos de esta solución de la rodilla del frente se muestra en la Figura 5. En este caso, se formaron 5 grupos, de los cuales 3 tienen solamente un sensor. Los otros 2 grupos tiene 9 y 8 sensores, respectivamente. Como se puede observar en la figura, los líderes de esta solución son aquellos sensores más cercanos a la base. Aunque no se muestra en este documento, para instancias con muchos más sensores (300 o 400) el sensor más cercano a la base también fue seleccionado como líder del grupo.

5.2. Relación de conflicto entre los objetivos

En esta sección se presenta un análisis para conocer la relación de conflicto entre cada par de los objetivos del problema de selección del líder. Para estimar el conflicto entre los objetivos utilizamos el coeficiente de correlación de Spearman. Este coeficiente mide la dependencia entre dos variables representadas por dos muestras de puntos (en nuestro caso, el conjunto de soluciones en el espacio objetivo). A diferencia del coeficiente de correlación de Pearson, la dependencia no necesariamente debe ser lineal, ya que solamente necesita describir una función monótona (creciente o decreciente). Los valores del coeficiente de Spearman caen en el intervalo $[-1, 1]$, donde los valores límite -1 y 1 significan que el par de variables siguen un orden monótono perfecto decreciente o creciente, respectivamente. En términos de conflicto, dos objetivos están en conflicto si el coeficiente de correlación es -1 , mientras que los objetivos se apoyan cuando el valor de correlación es 1 . Si tenemos un valor cercano a cero significa que los objetivos son independientes.

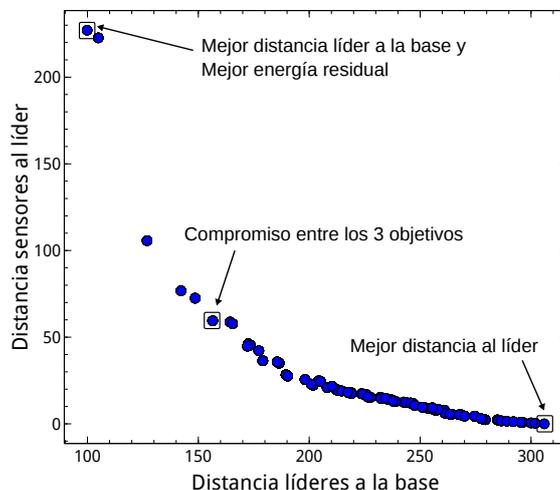


Fig. 2. Vista de 2 objetivos del frente de Pareto de una instancia con 20 sensores.

Para aplicar este coeficiente de correlación al problema de los líderes, utilizamos el NSGA-II para generar aproximaciones del frente de Pareto. Para obtener una buena aproximación del frente, para cada instancia, utilizamos una población de 2000 individuos durante 20000 generaciones. Es decir, un total de 4×10^7 evaluaciones de las tres funciones objetivo. Este número de evaluaciones significa que la porción estimada de espacio explorado por NSGA-II es de $4 \times 10^7 / 2^{100} = 3.1 \times 10^{-23}, 2.4 \times 10^{-53}, 1.9 \times 10^{-83}, 1.5 \times 10^{-113}$, para instancias de 100, 200, 300 y 400 sensores.

Para determinar de manera objetiva si los objetivos están correlacionados o son independientes utilizamos una prueba estadística (nivel de significación de 1%) en la cual la hipótesis alternativa es que el coeficiente de Spearman sea diferente de cero para cada par de valores objetivo.

En la Tabla 2 se muestran los coeficientes de correlación de Spearman para cada par de objetivos del problema usando 5 instancias variando el número de sensores: 20, 100, 200, 300 y 400 sensores. En todos los casos que aparecen en dicha tabla se aceptó la hipótesis de que el coeficiente es diferente de cero. En términos generales, podemos decir que según el coeficiente de Spearman, hay claro conflicto entre f_1 y f_2 sin importar el número de sensores, es decir, la distancia de los líderes a la base y la distancia de los miembros de grupo a su líder. Esto quiere decir que si queremos reducir la distancia de los líderes a la base, esto se hará a expensas de empeorar la distancia de los miembros de un grupo a su líder.

De igual manera, hay conflicto entre f_2 y f_3 , la distancia de los miembros de grupo a su líder y la energía utilizada. Esto quiere decir que si hacemos más compactos los grupos, entonces la energía utilizada por los líderes aumentará en la mayoría de los casos. Por otra parte, entre los objetivos f_1 y f_3 (distancia de

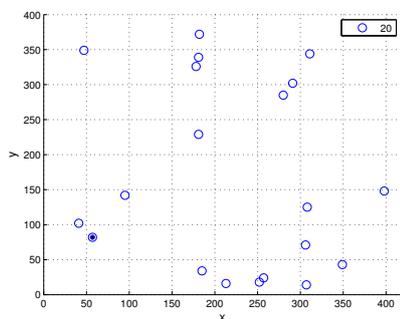


Fig. 3. Grupo único formado con 20 sensores para la solución con la mejor distancia de líderes a la base y la mejor energía residual. El círculo con un punto interior representa el único líder.

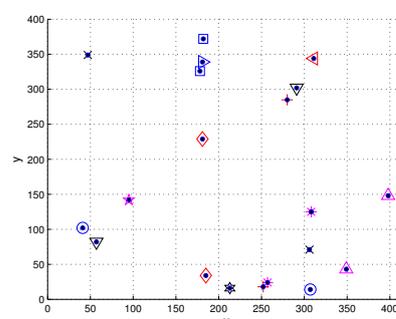


Fig. 4. Grupos formados con 20 sensores para la solución con la mejor distancia de los miembros a su líder. Los símbolos con un punto interior representan un líder.

los líderes a la base y la energía que utilizan) no hay conflicto en la mayoría de los casos.

Observando los resultados de la Tabla 2 para los casos de 300 y, principalmente, para 400 sensores, aunque aún la tendencia para f_1 y f_3 es de apoyo, y para f_2 y f_3 de conflicto, el valor absoluto de ambos valores no es tan grande como para los otros casos. La explicación de estos valores es clara al observar a detalle el frente de Pareto de f_1 y f_3 que se muestra en la Fig. 6. En la mayoría de las configuraciones de grupos formados, al disminuir la distancia de los líderes a la base también se reduce la energía que utilizan. No obstante, al acercarse a una distancia aproximada a 35m o menor, entonces la relación de dependencia cambia, es decir, se establece una relación de conflicto entre estos 2 objetivos. Esto explica por qué el coeficiente de Spearman aunque es positivo está más cercano a cero.

En términos de la red sensores este compromiso ocurre por la siguiente razón. La energía utilizada está determinada principalmente por el consumo necesario para enviar mensajes a la base (dependiente de la distancia). Sin embargo, también entra en juego la energía utilizada para recibir mensajes de los miembros del grupo (dependiente solamente del número de miembros). Así, la distancia a la base es prioridad para ahorrar energía, pero en cierto punto este ahorro ya no es suficiente para contrarrestar el gasto al recibir mensajes. Para nuestra instancia de 400 sensores distribuidos en un área de 400m \times 400m, este umbral de de 35m aproximadamente. Para otras instancias conocer el frente de Pareto sería útil para saber la distancia mínima para elegir los líderes sin aumentar la energía utilizada.

En la Fig. 7 se muestran la proyección del frente de Pareto con 400 sensores en los planos f_1 - f_3 y f_2 - f_3 . Nuevamente se resalta la zona de conflicto para f_1 - f_3 , pero es importante notar que también para la distancia de los miembros

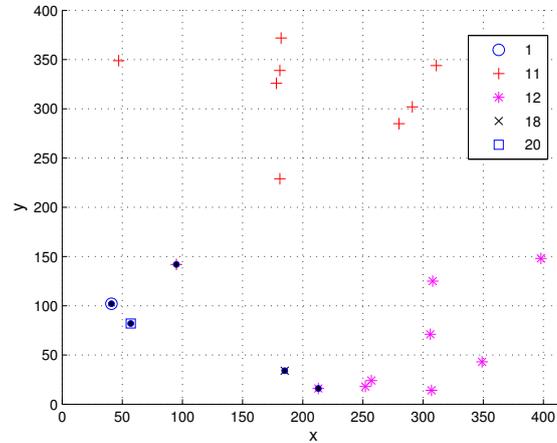


Fig. 5. Cinco grupos formados con 20 sensores para una solución que representa un compromiso intermedio entre los 3 objetivos. Los símbolos con un punto interior representan un líder.

de grupo a su líder y la energía utilizada, en la mayoría de los casos hay una relación de conflicto (de 0 a 50 metros aprox.). Sin embargo, más allá de 50m hay una relación de apoyo entre los objetivos, lo cual explica el valor del coeficiente de correlación negativo pero cercano a cero para 300 y 400 sensores. Por cuestiones de espacio no se muestra el frente de Pareto para 20 y 100 sensores, pero para estos 2 casos, los frentes no tienen regiones mixtas de apoyo y conflicto. Es decir, para f_1 - f_3 toda la curva es de apoyo, mientras que para f_2 - f_3 solamente hay relación de conflicto.

Tabla 2. Coeficientes de correlación de Spearman para estimar el conflicto entre objetivos.

# sensores	f_1 vs f_2	f_1 vs f_3	f_2 vs f_3
20	-0.9993	0.9990	-0.9997
100	-0.9998	0.9685	-0.9687
200	-1.0000	0.9996	-0.9996
300	-1.0000	0.9975	-0.9976
400	-0.9203	0.2907	-0.4281

6. Conclusiones

En este documento se presentó un análisis de tres de los objetivos que se pueden encontrar comúnmente en la literatura minimizar la distancia entre

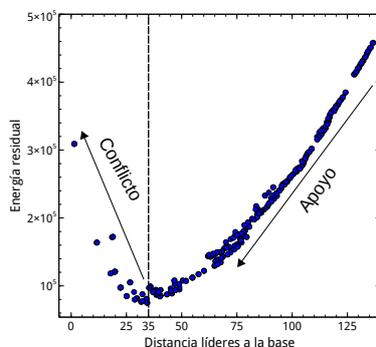


Fig. 6. Ilustración del conflicto entre distancia de los líderes a la base y la energía utilizada a partir de cierta distancia.

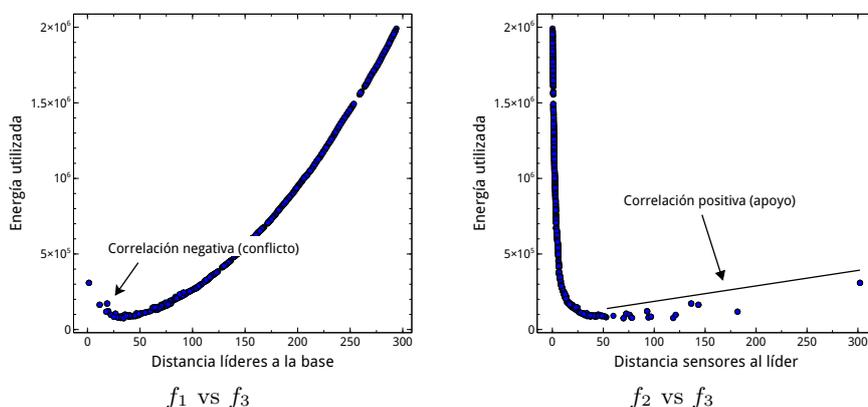


Fig. 7. Frente de Pareto obtenido para una instancia con 400 sensores.

miembros del grupo, minimizar la distancia de los líderes a la base y maximizar la energía residual de los líderes. Para ello, realizamos un estudio experimental utilizando el bien conocido algoritmo NSGA-II y donde analizamos la formación de grupos de sensores y la relación de conflicto entre los objetivos.

Respecto al análisis de conflicto entre los objetivos encontramos que bajo ciertas condiciones hay conflicto entre cada par de objetivos del problema de selección del líder. En particular, hay claro conflicto entre los dos pares de objetivos siguientes: distancia de los líderes a la base y distancia de los sensores a su líder; y distancia de los sensores a su líder y energía residual. Por otra parte, hay una relación mixta entre los objetivos distancia de los líderes a la base y energía residual. De una distancia entre 0 y 35 m (esta distancia depende del número de sensores y área en la que se distribuyen los sensores) hay una relación de conflicto, mientras que a una distancia mayor hay relación de apoyo. Esta distancia de umbral donde cambia la relación entre objetivos puede conocer hasta encontrar el frente de Pareto de cada instancia. En la práctica, los resultados

en cuanto a conflicto significan que los tres objetivos necesitan ser optimizados simultáneamente para optimizarlos ya que ningún objetivo es redundante.

En cuanto al tipo de grupos de sensores de la soluciones óptimos se observó que el líder de cada grupo siempre es el sensor más cercano a la base. De igual manera, la solución para minimizar la energía residual es una configuración con un solo grupo. Sin embargo, es importante notar que esta solución solamente toma en cuenta una ronda de comunicación donde no hay cambio de líderes.

Finalmente, como trabajo futuro extenderemos este estudio con más objetivos representativos de otras características de las redes; además consideraremos el escenario donde hay cambio de líderes y donde deberían reconfigurarse los *clusters*. Asimismo, se buscará la incorporación de los algoritmos evolutivos en un simulador de redes como NS-2 o WSNNet.

Referencias

1. Abo-Zahhad, M., Ahmed, S.M., Sabor, N., Sasaki, S.: A New Energy-Efficient Adaptive Clustering Protocol Based on Genetic Algorithm for Improving the Lifetime and the Stable Period of Wireless Sensor Networks. *International Journal of Energy, Information and Communications* (2014)
2. Afsar, M.M., Tayarani-N, M.H.: Clustering in sensor networks: A literature survey. *Journal of Network and Computer Applications* 46, 198–226 (Nov 2014)
3. Anastasi, G., Conti, M., Di Francesco, M., Passarella, A.: Energy conservation in wireless sensor networks: A survey. *Ad Hoc Networks* 7(3), 537–568 (2009)
4. Brockhoff, D., Zitzler, E.: Are All Objectives Necessary? On Dimensionality Reduction in Evolutionary Multiobjective Optimization. In: *Parallel Problem Solving from Nature - PPSN IX*, pp. 533–542 (2006)
5. Carlsson, C., Fullér, R.: Interdependence in fuzzy multiple objective programming. *Fuzzy Sets and Systems* 65(1), 19–30 (1994)
6. da Cunha, A.B., da Silva, D.C.: Behavioral Model of Alkaline Batteries for Wireless Sensor Networks. *IEEE Latin America Transactions* 10(1), 1295–1304 (Jan 2012)
7. Deb, K., Pratap, A., Agarwal, S., Meyarivan, T.: A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* 6(2), 182–197 (April 2002)
8. Elhabyan, R., Yagoub, M.C.: PSO-HC: Particle Swarm Optimization Protocol for Hierarchical Clustering in Wireless Sensor Networks. In: *Proceedings of CollaborateCom*. pp. 417–424. Miami, FL, USA (Oct 2014)
9. Iqbal, M., Naeem, M., Anpalagan, A., Qadri, N., Imran, M.: Multi-objective optimization in sensor networks: Optimization classification, applications and solution approaches. *Computer Networks* 99, 134–161 (Apr 2016)
10. Kheireddine, M., Abdellatif, R., Ferrari, G.: Genetic Centralized Dynamic Clustering in Wireless Sensor Networks, pp. 503–511 (2015)
11. Miranda, K.: Conexión total: Internet de las Cosas. ¿Cómo ves? *Revista de Divulgación de la Ciencia de la UNAM* 15(174), 22–25 (May 2013)
12. Miranda, K., Ramos, V.: Improving data aggregation in Wireless Sensor Networks with time series estimation. *IEEE Lat. Am. Trans.* 14(5), 2425–2432 (May 2016)
13. Slavik, M., Mahgoub, I., Badi, A., Ilyas, M.: Analytical model of energy consumption in hierarchical wireless sensor networks. In: *7th International Symposium on High-capacity Optical Networks and Enabling Technologies*. pp. 84–90 (Dec 2010)

14. Tubaishat, M., Madria, S.: Sensor networks: an overview. *IEEE Potentials* 22(2), 20–23 (Apr 2003)